

The Effects of Ignoring Level Shifts on Systems Cointegration Tests

by

Carsten Trenkler*
Sonderforschungsbereich 373
Humboldt-University Berlin
Spandauer Str. 1
D-10178 Berlin
Germany
Tel.: +49-30-2093-5705
Fax.: +49-30-2093-5712
email: trenkler@wiwi.hu-berlin.de

Abstract

In this paper I analyse the effects of ignoring level shifts in the data generating process (DGP) on systems cointegration tests that do not accommodate level shifts. I consider two groups of Likelihood Ratio tests based on procedures suggested by Johansen (1988, 1995) and Saikkonen & Lütkepohl (2000b). The Monte Carlo analysis reveals that ignoring level shifts reduces the tests' sizes to zero and causes an important drop in the small sample power for increasing shift magnitudes. These observations are also reflected in two empirical applications in such a way that the tests find a cointegrating rank smaller than one suggested by procedures which accommodate the shifts.

Keywords: Systems cointegration tests, Level shifts, Monte Carlo study

JEL classification: C32, C15

*This paper is adapted from a chapter of my PhD thesis. I would like to thank Christian Müller, Ralf Brüggemann, and Helmut Lütkepohl for many helpful comments and suggestions on an earlier version of the paper and the corresponding chapter of my thesis. Moreover, I am grateful to the Deutsche Forschungsgemeinschaft, SFB 373, for financial support.

1 Introduction

After the introduction of the cointegration concept by Granger (1981) and Engle & Granger (1987) it is now common practice to test for the cointegrating rank of a system of economic time series because it has an important impact on the further modelling strategy. Other inference tools, for example causality tests, impulse-response or forecasting analysis, depend on the cointegrating rank, i.e. on the number of cointegration relations. However, when analysing economic time series the researcher is often faced with structural breaks that effect these time series. Examples are German macroeconomic time series that exhibit a level shift due to the German unification or time series describing the development of macroeconomic variables of Central and Eastern European Countries which have undergone a transition process from a centrally planned to a market economy.

The question arises whether the occurrence of structural changes has an influence on the inference about the cointegrating rank.¹ This question refers to the decision problem of whether one may use standard cointegration tests which ignore structural changes in order to avoid possible power losses and size distortions caused by modelling structural shifts or whether it is recommended to use cointegration tests that take such breaks into account. The latter strategy should be pursued if the presence of structural breaks distorts the inference about the cointegrating rank, i.e. if the standard tests would no longer be a valid inference tool. Such a strategy would be necessary because an incorrect cointegrating rank can lead to a wrong economic interpretation of the behaviour of the system analysed and, furthermore, it may have a negative impact on other inference tools some of which are mentioned above.

Within the univariate framework the presence of structural changes is regarded as very important when testing for a unit root. Perron (1989) has shown for the Augmented Dickey Fuller test that the null hypothesis of a unit root is clearly underrejected if a break in the linear trend term or shift in the level of a time series occur. Maddala & Kim (1998) and Yin & Maddala (1997) have found that depending on the way the structural break is modelled overrejection of the null hypothesis of a unit root can also occur. Hence, no matter which certain kind of a structural break is modelled the overall message is that the presence of structural changes may distort the inference about a unit root.

Thus, if the presence of changes distorts the inference in the univariate framework we may expect similar effects in the multivariate framework when testing for the number of cointegration relations. This expectation is further supported by evidence from small sample studies conducted by Gregory & Hansen (1996) and Gregory, Nason & Watt

¹I regard the terms *breaks* and *changes* as equivalent.

(1996) who have found that breaks in the cointegration relation reduces the power of single-equation cointegration tests.

In line with the foregoing discussion this paper analyses whether structural changes distort the inference of systems cointegration tests ignoring such breaks. To be more precise, I consider a shift in the level of the time series as a special kind of structural breaks. This analysis is conducted on the basis of a Monte Carlo study and reviewing two empirical examples. I consider two groups of Likelihood Ratio tests. The first group of tests has been suggested by Johansen (1991) and Johansen (1992, 1994, 1995). The former test allows for a mean term and the latter for a linear trend term as well.² Both tests are based on a procedure proposed by Johansen (1988). Therefore, these proposals are referred to as Johansen tests which have become the most popular systems cointegration tests in applied work. The second group of tests has been proposed by Saikkonen & Lütkepohl (2000b). These authors also introduce two test versions allowing for a mean term and a linear trend. They are denoted as Saikkonen-Lütkepohl tests. Note again that the aforementioned procedures do not take account of any kind of structural breaks.

The focus on level shifts is regarded as a starting point for a more general analysis of the effects of structural changes. Moreover, it is expected that if the inference of the tests is distorted by the presence of such a simple break type the properties of the tests are even worse if more complicated breaks are considered.

The paper is organized as follows. In the next section I present the theoretical framework and the cointegration tests. Findings of previous related simulation studies relevant for my analysis are reviewed in Section 3. The tests' small sample size and power properties based on my Monte Carlo simulations are described in Section 4. Section 5 contains the empirical examples referring to the Polish crawling peg system and a German money demand system respectively. In both models some of the time series exhibit a level shift. Summarizing and concluding remarks are given in Section 6.

2 Theoretical Framework

In general, the considerations are based on a n -dimensional times series $y_t = (y_{1t}, \dots, y_{nt})'$ ($t = 1, \dots, T$) which is generated by

$$y_t = \mu_0 + \mu_1 t + \delta d_t + x_t, \quad t = 1, 2, \dots, \quad (2.1)$$

where μ_i ($i = 0, 1$) and δ are unknown $(n \times 1)$ parameter vectors. Hence, the deterministic part consist of a mean term μ_0 , a linear trend term $\mu_1 t$ and a level shift term δd_t where

²The test due to Johansen (1991) has also been considered by Johansen & Juselius (1990).

d_t is a shift dummy variable defined as

$$d_t = \begin{cases} 0, & t < T_1 \\ 1, & t \geq T_1 \end{cases}. \quad (2.2)$$

This shift dummy allows to take into account a sudden change in the mean or the level of the process. To simplify the exposition I consider only one level shift. It is possible to generalize the framework by including further shift dummy or impulse dummy variables.

The term x_t is an unobservable stochastic error process which is assumed to follow a vector autoregressive process of order p (VAR(p)):

$$x_t = A_1 x_{t-1} + \cdots + A_p x_{t-p} + \varepsilon_t, \quad (2.3)$$

where A_j , ($j = 1, \dots, p-1$) are $(n \times n)$ coefficient matrices and $\varepsilon_t \sim N(0, \Omega)$. The assumption of normally distributed error terms is made to simplify the establishment of a common framework for comparing the different test procedures (compare e.g. Trenkler 2002). The initial values x_{-p+1}, \dots, x_0 are assumed to be from some fixed probability distribution which does not depend on the sample size. The usual vector error correction model (VECM) form is

$$\Delta x_t = \Pi x_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta x_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots, T, \quad (2.4)$$

where $\Pi = -(I_n - A_1 - \cdots - A_p)$ and $\Gamma_j = -(A_{j+1} + \cdots + A_p)$ ($j = 1, \dots, p-1$) are $(n \times n)$ matrices. It is assumed that x_t is at most $I(1)$ and cointegrated with a rank r . This implies that y_t is also at most $I(1)$ and cointegrated with a rank r . Moreover, it follows that the matrix Π can be written as

$$\Pi = \alpha \beta', \quad (2.5)$$

where α and β are $(n \times r)$ matrices of full column rank. When testing for the cointegrating rank one tests for the rank of the matrix Π . There exist two variants of the considered systems tests with respect to the rank specified under the alternative hypothesis. The maximum eigenvalue variant tests against a cointegrating rank that is larger by one than the rank specified under the null hypothesis. In contrast, the trace variant tests against a more general alternative which states that the rank is larger than the one in the null hypothesis. For a situation of no level shifts Lütkepohl, Saikkonen & Trenkler (2001) have compared these two test versions for various Johansen-type and Saikkonen-Lütkepohl proposals by means of a local power and a small sample analysis. They have found that

the properties of both versions are similar in general. Therefore I only consider the trace test version assuming that the findings also hold in a set-up with level shifts. To be specific, the pair of hypotheses

$$H_0(r_0) : \text{rk}(\Pi) = r_0 \quad \text{vs.} \quad H_1(r_0) : \text{rk}(\Pi) > r_0. \quad (2.6)$$

is tested. In fact, limited simulation evidence has shown that both test versions perform similarly if level shifts are ignored.

To present the tests it is useful to write (2.1) in a VECM representation using (2.4):

$$\begin{aligned} \Delta y_t &= \nu_0 + \nu_1 t + \Pi y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=0}^{p-1} \gamma_j^* \Delta d_{t-j} + \delta_\Pi d_t + \varepsilon_t, \\ t &= p+1, p+2, \dots, \end{aligned} \quad (2.7)$$

where $\nu_0 = -\Pi\mu_0 + (I_n + \Pi - \sum_{j=1}^{p-1} \Gamma_j)\mu_1$, $\nu_1 = -\Pi\mu_1$, $\delta_\Pi = -\Pi\delta$ and

$$\gamma_j^* = \begin{cases} (I_n + \Pi)\delta, & j = 0 \\ -\Gamma_j\delta, & j = 1, \dots, p-1 \end{cases}. \quad (2.8)$$

Note that Δd_{t-j} is an impulse dummy which has the value one at $t = T_1 + j$ and is zero elsewhere.

As we know the structure of ν_1 and δ_Π we may restrict the linear trend and the lagged shift dummy to the cointegration relations and write (2.7) as

$$\begin{aligned} \Delta y_t &= \nu + \alpha[\tau(t-1) + \phi d_{t-1} + \beta' y_{t-1}] + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=0}^{p-1} \gamma_j \Delta d_{t-j} + \varepsilon_t \\ &= \nu + \Pi^+ y_{t-1}^+ + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=0}^{p-1} \gamma_j \Delta d_{t-j} + \varepsilon_t, \quad t = p+1, p+2, \dots, \end{aligned} \quad (2.9)$$

with $\nu = \nu_0 + \nu_1$, $\tau = -\beta'\mu_1$, $\phi = -\beta'\delta$, $\Pi^+ = \alpha[\tau : \phi : \beta']$, $y_{t-1}^+ = [(t-1) : d_{t-1} : y'_{t-1}]'$ and

$$\gamma_j = \begin{cases} \delta, & j = 0 \\ -\Gamma_j\delta, & j = 1, \dots, p-1 \end{cases}. \quad (2.10)$$

The expression $\Pi^+ y_{t-1}^+$ is the error correction term which comprises the cointegration relations with the restricted deterministic terms.

However, in the simulation study tests are applied that ignore level shifts. That means, the tests are actually based on (2.1) with $\delta = (0, \dots, 0)'$. Hence, $\gamma_j = 0$ and $\delta_\Pi = 0$. Then, we can write the restricted VECM (2.9) of y_t as

$$\begin{aligned}
\Delta y_t &= \nu + \alpha[\tau(t-1) + \beta' y_{t-1}] + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t \\
&= \nu + \Pi^+ y_{t-1}^+ + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t, \quad t = p+1, p+2, \dots,
\end{aligned} \tag{2.11}$$

with $\nu = \nu_0 + \nu_1$, $\Pi^+ = \alpha[\tau : \beta']$ and $y_{t-1}^+ = [(t-1) : y'_{t-1}]'$.

If we assume $\mu_1 = 0$, i.e. making no allowance for a linear trend, then $\nu_0 = -\Pi\mu_0$, $\nu_1 = 0$, and hence $\nu = \nu_0$. So, (2.11) can be written as

$$\begin{aligned}
\Delta y_t &= \alpha[\tau_0 + \beta' y_{t-1}] + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t \\
&= \Pi^0 y_{t-1}^+ + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t, \quad t = p+1, p+2, \dots,
\end{aligned} \tag{2.12}$$

with $\tau_0 = -\beta'\mu_0$, $\Pi^0 = \alpha[\tau_0 : \beta']$ and $y_{t-1}^0 = [1 : y'_{t-1}]'$.

The Johansen tests are performed by a maximum likelihood estimation of the parameters in (2.11) and (2.12) respectively leading to a generalized eigenvalue problem. For $\mu_1 \neq 0$ the procedures works as follows. Δy_t and y_{t-1}^+ are regressed on $(1, \Delta y_{t-1}, \dots, \Delta y_{t-p})$ to obtain the residuals R_{0t} and R_{1t} , respectively. Next, consider $S_{ij} = T^{-1} \sum_{t=1}^T R_{it} R'_{jt}$ ($i, j = 0, 1$). Solving $\det(\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}) = 0$ we obtain the ordered generalized eigenvalues $\lambda_1 \geq \dots \geq \lambda_n$. Then the LR test statistic for the pair of hypothesis in (2.6) is

$$LR_{Joh}(r_0) = -T \sum_{j=r_0+1}^n \log(1 - \lambda_j). \tag{2.13}$$

The test statistic for the case $\mu_1 = 0$ can be derived accordingly by regressing Δy_t and y_{t-1}^0 on $(\Delta y_{t-1}, \dots, \Delta y_{t-p})$ and proceed otherwise as above. The ordered generalized eigenvalues are denoted by $\lambda_1^0 \geq \dots \geq \lambda_n^0$ and the test statistic is

$$LR_{Joh}^0(r_0) = -T \sum_{j=r_0+1}^n \log(1 - \lambda_j^0). \tag{2.14}$$

The limiting distributions of $LR_{Joh}(r_0)$ and $LR_{Joh}^0(r_0)$ are nonstandard and functions of $(n - r_0)$ -dimensional standard Brownian motions. Interestingly, the limiting distributions depend on $(n - r_0)$, the number of common stochastic trends under the null hypothesis, but not on n and r_0 separately. Furthermore, they are not affected by the actual values of

the mean and trend parameter μ_0 and μ_1 respectively. Critical values and further details can be found, for example, in Johansen (1995, Chapter 15).

As suggested by Johansen (1991) and Perron & Campbell (1993) it is also possible to formulate test statistics based on unrestricted VECM models analogous to (2.7) by setting the shift parameter to zero. However, in (2.7) the terms ν_0 and $\nu_1 t$ may generate a quadratic trend, what is ruled out by the assumptions underlying the DGP (2.1). Therefore, I apply the tests that impose the proper restrictions on the deterministic terms by confining these terms to the long-run relationship.

Finally, there exists a Johansen test version which assumes that the variables have a linear trend but the trend is orthogonal to the space spanned by the cointegrating vectors (see Johansen 1991, Johansen & Juselius 1990). The latter assumption means that the trend parameter vector μ_1 is orthogonal to the cointegration matrix β' , i.e. $\beta'\mu_1 = 0$. Despite the popularity of this test version in applied work we have to be aware of its disadvantages. First of all, it is questionable whether the assumption of an orthogonal trend can be verified easily for real data sets. This is important, however, because if the trend is not orthogonal to the cointegration space the test version is inappropriate. Moreover, the limiting distribution under the null hypothesis depends on whether the linear trend is present in the data or not. Johansen (1995, Chapters 11 & 12) gives a detailed discussion on that issue. In fact, if the linear trend is absent the limiting distribution has broader tails. Hence, using critical values which are derived from the distribution based on the existence of a trend may result in overrejection of the null hypothesis. Indeed, the excessive size distortion turned out to be important in small samples if data without a linear trend are used (compare Hubrich, Lütkepohl & Saikkonen 2001). For all these reasons, the test version assuming an orthogonal trend is not considered in the following.

The test proposal made by Saikkonen & Lütkepohl (2000a) allowing for a linear trend refers to the process (2.1) with $\delta = (0, \dots, 0)$. Their idea is to estimate the deterministic terms in a first step by a feasible generalized least squares (GLS) procedure and adjust y_t by these estimated terms. In a second step, a Johansen-type test is performed on the adjusted time series. More precisely, defining $A(L) = I_n - A_1 L - \dots - A_p L^p = I_n \Delta - \alpha \beta' L - \Gamma_1 \Delta L - \dots - \Gamma_{p-1} \Delta L^{p-1}$ and

$$a_{0t} = \begin{cases} 1 & \text{for } t \geq 1 \\ 0 & \text{for } t \leq 0 \end{cases} \quad \text{and} \quad a_{1t} = \begin{cases} t & \text{for } t \geq 1 \\ 0 & \text{for } t \leq 0 \end{cases}$$

we obtain by multiplying (2.1) by $A(L)$

$$A(L)y_t = H_{0t}\mu_0 + H_{1t}\mu_1 + \varepsilon_t, \quad t = 1, \dots, T, \quad (2.15)$$

where $H_{it} = A(L)a_{it}$ ($i = 0, 1$), and $\varepsilon_t = A(L)x_t$. Furthermore, defining Q such that $QQ' = \Omega^{-1}$ and multiply (2.15) by Q' , we have

$$Q'A(L)y_t = Q'H_{0t}\mu_0 + Q'H_{1t}\mu_1 + \eta_t, \quad t = 1, \dots, T, \quad (2.16)$$

where $\eta_t = Q'\varepsilon_t$. Hence, the error term η_t has a zero mean and a unit covariance matrix as it is required for a GLS transformation. To make the GLS estimation feasible Saikkonen & Lütkepohl (2000a) propose to use the ML estimators $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\Gamma}_j$ ($j = 1, \dots, p-1$) and $\tilde{\Omega}$ obtained by a reduced rank (RR) regression of (2.11) applying the rank r_0 which is specified under the null hypothesis of the cointegration test. Using these estimators one can compute the estimators \hat{Q} , $\hat{A}(L)$, and \hat{H}_{it} ($i = 0, 1$). Thus, feasible GLS estimators $\hat{\mu}_0$, $\hat{\mu}_1$, and $\hat{\delta}$ of μ_0 , μ_1 , and δ are obtained by a multivariate least squares (LS) estimation of the model

$$\hat{Q}'\hat{A}(L)y_t = \hat{Q}'\hat{H}_{0t}\mu_0 + \hat{Q}'\hat{H}_{1t}\mu_1 + \eta_t^*, \quad t = 1, \dots, T. \quad (2.17)$$

Since the cointegrating rank under the null hypothesis r_0 is used in estimating the parameters in (2.9), the deterministic terms are also estimated under $H_0(r_0)$. Having estimated the deterministic terms one can adjust y_t by these terms and form a sample analog $\hat{x}_t = y_t - \hat{\mu}_0 - \hat{\mu}_1 t$ of x_t . Then, as already mentioned, a LR-type test is performed on \hat{x}_t . Since \hat{x}_t is adjusted by the deterministic terms a test version without any deterministic terms like in Johansen (1988) is applied. However, the generalized eigenvalue problem as formulated in Saikkonen & Lütkepohl (2000b) has a different structure. Consider first the feasible VECM of the sample analog \hat{x}_t

$$\Delta\hat{x}_t = \Pi\hat{x}_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta\hat{x}_{t-j} + e_t, \quad t = p+1, \dots, T \quad (2.18)$$

where e_t is an error term. Let $\hat{\Pi}$ be the LS estimator of Π in (2.18) and $\hat{\Omega}$ be the corresponding estimated covariance matrix of the error terms. Furthermore, define $\hat{S}_{11} = T^{-1} \sum_{t=1}^T \hat{R}_{1t} \hat{R}'_{1t}$ where \hat{R}_{1t} are the residuals obtained from regressing \hat{x}_{t-1} on $(\Delta\hat{x}'_{t-1}, \dots, \Delta\hat{x}'_{t-p})'$. Then, the LR-type test statistic is derived by solving the generalized eigenvalue problem $\det(\hat{\Pi}\hat{S}_{11}\hat{\Pi} - \lambda\hat{\Omega}) = 0$. Using the resulting eigenvalues $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_n$ the test statistic for the pair of hypotheses in (2.6) is

$$LR_{S\&L}(r_0) = T \sum_{j=r_0+1}^n \log(1 + \hat{\lambda}_j). \quad (2.19)$$

Note that the two kinds of generalized eigenvalue problems described here and applied for the Johansen tests can be transformed into each other by an appropriate redefinition

of the respective eigenvalues. Hence, the LR statistics based on the two different sets of eigenvalues are identical apart from minor numerical differences.

If we assume that the linear trend is absent, we set $\mu_1 = 0$ in the above procedure. Accordingly, the ML estimators are derived from (2.12) and a LR-type test is applied on the series $\hat{x}_t = y_t - \hat{\mu}_0$. The resulting test statistic is named $LR_{S\&L}^0(r_0)$.

The limiting distributions under the null hypothesis and further details of the test procedures are given in Saikkonen & Lütkepohl (2000b). The limiting distribution of $LR_{S\&L}(r_0)$ is a function of $(n - r_0)$ -dimensional Brownian bridges. Replacing these Brownian bridges by $(n - r_0)$ -dimensional standard Brownian motions gives the corresponding limiting distribution for $LR_{S\&L}^0(r_0)$. Like the null distributions of the Johansen tests those of $LR_{S\&L}(r_0)$ and $LR_{S\&L}^0(r_0)$ depend on $n - r_0$, not on n and r_0 separately and they are independent of the actual values of μ_0 and, regarding $LR_{S\&L}(r_0)$, independent of μ_1 . Critical values are tabulated in Lütkepohl & Saikkonen (2000).

3 Previous Simulation Studies

So far, we know from a number of studies dealing with single-equation cointegration tests in the presence of structural breaks that neglecting breaks may result in distorted inference. Campos, Ericsson & Hendry (1996) analyse within a bivariate system the effects of a level shift in the stationary marginal process on different cointegration tests performing a Monte Carlo study. The tests' size is only influenced a little. Furthermore, tests based on estimated error correction models are generally more powerful than the two-step testing procedure suggested by Engle & Granger (1987). The latter procedure is often called the Augmented Dickey-Fuller (ADF) test for cointegration. The findings of Campos et al. (1996) contrasts with results of Gregory et al. (1996) and Gregory & Hansen (1996) for the ADF test for cointegration. Gregory et al. (1996) analyse a break in the coefficients of the cointegrating vector in bivariate systems. Gregory & Hansen (1996) use a bivariate system like in Engle & Granger (1987) to study level shifts and trend breaks in the cointegrating relation. They all observe a clear underrejection of the null of no cointegration in their Monte Carlo study. For example, in the experiment of Gregory & Hansen (1996) the ADF test for cointegration with a mean term loses about two third of its original power of 80% if a level shift with a magnitude of three times the standard deviation of the error terms is introduced to the DGP. The power of the test variant allowing for a linear trend decreases by about one third of its original level of 70%. In case of a trend break the power reduction is even more severe.

Performing different Monte Carlo experiments Doornik, Hendry & Nielsen (1998) and Inoue (1999) have shown that this observation may carry over to Johansen-type tests when breaks are not taken into account properly. Inoue (1999) uses exactly the same data generating process as Gregory & Hansen (1996). He analyses the test versions suggested by Johansen (1991) and Perron & Campbell (1993) which I do not consider. Using a level shift of the same magnitude like in Gregory & Hansen (1996), Inoue (1999) reports a reduction in power from about 95% to approximately 70% when a level shift is ignored. Again, a trend break has an even more distortionary effect.

In one of their experiments that is relevant for the analysis presented here Doornik et al. (1998) add an unrestricted impulse dummy to two components of a VECM process like (2.11) with two cointegration relationships and a restricted linear trend. The dummy variables have a magnitude of five and three times the standard deviation of the error terms respectively. The Johansen test without any dummy variables has still reasonable small sample size and power.³ Unfortunately, the authors do not present results for the situation of no dummy variables in the DGP. Therefore, we cannot assess the change in the test's size and power caused by the introduction of a dummy variable. The experiment was repeated with dummies of larger magnitudes approximately equal to fifty and thirty times the standard deviation of the error terms respectively. In this case one can observe a slight reduction in the size and a fall in the small sample power from 50% to 35% compared to the case of small dummies if a sample size of $T = 100$ is considered. Thus, it seems that ignoring unrestricted impulse dummies causes only minor problems.

Depending on the true parameters, however, an unrestricted impulse dummy variable within a VECM may have, like other unrestricted deterministic terms, an ambiguous relationship to corresponding variables in the level representation of the time series. An unrestricted impulse dummy may have been generated by a shift in the level of the time series but it can be produced by simple outliers in the levels as well. Note that the DGPs used by Gregory & Hansen (1996) and Inoue (1999) refer to breaks in the cointegration relation. In contrast to all these studies I specify a shift in the levels of the individual time series and separate the shift from the stochastic part of the process. The advantage of this way of modelling is that the different effects of the shift in the VECM representation can be clearly seen. From (2.9) we find that the level shift in (2.1) produces a shift in the cointegration relations and a set of lagged unrestricted impulse dummy variables in a VECM. In fact, in the simulation study I use a DGP that enables us to easily distinguish the situation where a shift occurs in the cointegration relations from the case where only

³Doornik et al. (1998) just consider the test variant allowing for a trend (LR_{Joh}).

unrestricted impulse dummies are generated in the corresponding VECM. The ability to distinguish these two situations will help to relate my findings with the results of the other studies cited before.

The framework used in Doornik et al. (1998) is the one that is most closely related to my study. I extend, however, the analysis of Doornik et al. (1998) with respect to certain aspects which refer to the dimensionality of the DGPs, the correlation between the error terms, the sample length, the cointegrating rank, the timing and the magnitude of the level shifts. Moreover, I consider cointegration tests which allow for a linear trend as well as those which assume no trend.

So far we have just discussed small sample results. In fact, there are no results about the asymptotic behaviour of the standard cointegration tests available if structural breaks in the DGP are ignored. We have results, however, regarding the ADF unit root test in case of neglecting a level shift. Yin & Maddala (1997) state that for a fixed size of the level shift the null distribution of the test does not change but they expect an additional small sample bias due to the level shift. In contrast, for a shift magnitude that increases proportional to the square root of the sample size or even faster, Doornik et al. (1998) have shown that the null distribution is affected by the shift. These findings correspond to results in Perron (1989, 1990) with respect to the test's asymptotic behaviour under the alternative hypothesis. Under the alternative, the autoregressive parameter of the first lag in the corresponding univariate time series model is overestimated. But only for a shift magnitude increasing at least proportional to the square root of the sample size the estimated value asymptotically approaches one, which means that the process is regarded as nonstationary. Hence, in this case the ADF test cannot reject a wrong null hypothesis asymptotically. On the other hand, for a fixed shift size there is some chance that a wrong null hypothesis is rejected. Nevertheless, one should be cautious when comparing these asymptotic results for the univariate case with the outcome of the Monte Carlo analysis presented in the next section since the situation for cointegrated systems with a nonzero cointegrating rank can be much more complicated.

4 Monte Carlo Simulations

4.1 The DGP and Simulation Details

The simulations are based on the following DGP:

$$y_t = \delta d_t + x_t, \tag{4.1}$$

where δ is a $(n \times 1)$ vector with $\delta = (\delta_1, \dots, \delta_n)'$ measuring the magnitude of the level shift in the respective components of y_t , d_t is a shift dummy variable as defined in (2.2). Hence, I only consider a single shift at one point in time. This is done in order to simplify the analysis. It is expected that the existence of more than one level shift worsens the situation for the standard cointegration tests. The component x_t is a VAR(1) process that was already used and analysed by Toda (1994, 1995):

$$x_t = A_1 x_{t-1} + \varepsilon_t = \begin{bmatrix} \Psi & 0 \\ 0 & I_{n-r} \end{bmatrix} x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid } N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_r & \Theta \\ \Theta' & I_{n-r} \end{bmatrix}\right), \quad (4.2)$$

where $\Psi = \text{diag}(\psi_1, \dots, \psi_r)$ is an $(r \times r)$ diagonal matrix and Θ is a $(r \times (n-r))$ matrix describing the correlation between the stationary and nonstationary components. Other VAR(1) processes of interest can be obtained from (4.2) by linear transformations which leave the analysed LR tests invariant (see Toda (1994, 1995)). Specifically, I will consider bivariate and four-dimensional processes. For instance, in the bivariate case, if $r = 0$, Ψ and Θ vanish and the process consists of two nonstationary components. If the cointegrating rank is 1, $\Psi = \psi_1$ with $|\psi_1| < 1$. In that case $\Theta = \theta$ is a scalar which represents the instantaneous correlation between the two components. In the simulations the parameter values of the constant and the linear trend are zero, i.e. $\mu_i = 0$ ($i = 0, 1$) throughout. This is done because the test results are invariant to specific parameter values of the deterministic components they allow for. Thereby I can use one DGP for all tests no matter which assumptions are made regarding a constant or a linear trend.

According to (2.7) and (2.9), (4.1) has the following VECM representation:

$$\begin{aligned} \Delta y_t &= \alpha[\phi d_{t-1} + \beta' y_{t-1}] + \delta \Delta d_t + \varepsilon_t \\ &= [\delta_\Pi d_{t-1} + \Pi y_{t-1}] + \delta \Delta d_t + \varepsilon_t = \Pi^d y_{t-1}^d + \delta \Delta d_t + \varepsilon_t, \end{aligned} \quad (4.3)$$

with $\delta_\Pi = \alpha\phi = -\alpha\beta'\delta = -\Pi\delta$, $\Pi = \alpha\beta' = -(I_n - A_1)$, $\Pi^d = [\delta_\Pi : \Pi]$ and $y_{t-1}^d = [d_{t-1} : y'_{t-1}]'$. Hence, for δ_Π we have

$$\delta_\Pi = (I_n - A_1)\delta = \begin{bmatrix} (I_r - \Psi) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta^s \\ \delta^n \end{bmatrix} = \begin{bmatrix} (I_r - \Psi)\delta^s \\ 0 \end{bmatrix}, \quad (4.4)$$

where $\delta^s = (\delta_1, \dots, \delta_r)$ and $\delta^n = (\delta_{r+1}, \dots, \delta_n)$ are $(r \times 1)$ and $((n-r) \times 1)$ vectors representing shifts in the r stationary and $(n-r)$ nonstationary components of the DGP respectively.

Obviously, it matters whether the shift occurs in a stationary or in a nonstationary component. A level shift in a nonstationary component leads only to an unrestricted impulse dummy (Δd_t) in the VECM representation (4.3) since the respective coefficient

in δ_{Π} is zero according to (4.4). This latter fact means that the parameter vector δ is orthogonal to the space spanned by the cointegrating vectors contained in β' . The vector δ has to be orthogonal to the cointegration space because the respective coefficients in $\delta_{\Pi} = -\Pi\delta = -\alpha\beta'\delta$ are zero although α , β and δ are nonzero parameter vectors or matrices.⁴ In contrast, a shift in a stationary component of y_t also results in a shift in the cointegration relations or in the error correction term $\Pi^d y_{t-1}^d$. Generalizing this observation I can state that a shift in the level of a time series only generates a shift in a cointegration relation if the corresponding parameter vector of the level shift is nonorthogonal to the cointegration space. This property is analogous to other deterministic terms as e.g. to a linear trend (compare the discussion in Section 2 regarding the Johansen test version with an orthogonal trend).

In line with the foregoing remarks, I will refer to the level shifts generating a shift in the cointegration relations as nonorthogonal shifts whereas shifts leading only to unrestricted impulse dummies in the VECM representation will be named as orthogonal level shifts. In this respect, we may interpret the unrestricted impulse dummies used in Doornik et al. (1998) as being generated by an orthogonal level shift. In the Monte Carlo simulations the relative importance of nonorthogonal and orthogonal level shifts is analysed. On the other hand, the unrestricted impulse dummies may still correspond to simple outliers in the level or VAR representation of the DGP. Such dummy variables, however, are not considered in the current framework. Accordingly, the level shift in the cointegration relation studied by Gregory & Hansen (1996) and Inoue (1999) may be interpreted as a nonorthogonal level shift. Nonetheless, the resulting lagged impulse dummies are ignored in this way of modelling.

The computations for the simulation study are performed by using programs written in GAUSS 3.2 for Windows. The RNDNS function with a fixed seed has been used to generate standard normally distributed random numbers. Samples of sizes 100 and 200 plus 50 presample values starting with initial values of zero are generated and the presample values are discarded afterwards. The rejection frequencies given in Tables 1 - 5 and Figures 1 - 10 are based on asymptotic critical values for a significance level of 5%.

The number of replications is 10000, i.e. I generate 10000 different sets of multivariate random numbers. Note that I produce these random numbers only once. Hence, for each set of parameter values I use the same random numbers. This also means that the test statistics are applied to the same time series for a given set of parameter values. Thus, the

⁴The matrices or vectors α and β are only nonzero if the cointegrating rank r is larger than zero. Accordingly, if $r = 0$ the cointegration space is a null space and, thus, only shifts orthogonal to the cointegration space exist.

rejection frequencies reported in the analysis are not independent. However, they can be compared directly. In order to evaluate the results it may be helpful to use the standard error of an estimator of a true rejection probability P which is $s_P = \sqrt{P(1-P)/10000}$ if 10000 replications are considered. Hence, for $P = 0.05$ we obtain $s_{0.05} = 0.0022$ and the corresponding two-standard-error confidence interval is $[0.0456, 0.0544]$. Furthermore, the tests were not performed sequentially in the simulations. Thus, the results for testing $H_0(1) : \text{rk}(\Pi) = 1$ are not conditioned on the outcome of the test of $H_0(0) : \text{rk}(\Pi) = 0$ etc..

Doornik et al. (1998) suggest to scale the shift magnitude in order to be able to evaluate the shift's impact in a meaningful way. For the univariate case they introduce the scaled parameter $\delta^* = \delta/(\sqrt{T}\sigma)$ where σ is the standard deviation of the error term in the respective DGP. This scaling formula is motivated by the asymptotic results of Doornik et al. (1998) for the ADF unit root test mentioned in Section 3.1. If the parameter δ is fixed and the sample size T increases, the scaled parameter δ^* approaches zero. In this case the level shift is asymptotically negligible. On the other hand if δ increases with a rate proportional to the square root of the sample size, δ^* remains constant and the level shift has an effect on the asymptotic distribution of the ADF test.

However, it is not obvious whether the scaling formula suggested by Doornik et al. (1998) should also be applied in the multivariate framework with respect to the cointegration tests. For example, the formula is not able to capture the correlation between different error terms of a multivariate process. Nevertheless, as I set $\sigma^2 = 1$ for all components of the DGPs one should take the scaling rule $\delta^* = \delta/\sqrt{T}$ into account when comparing the results for the sample sizes $T = 100$ and $T = 200$. In the present context, δ^* and δ refer to parameter vectors instead of single parameters. Consider, for example, a bivariate parameter vector $\delta = (1 \ 1)'$ regarding a sample size $T = 100$. Then, $\delta^* = (1 \ 1)'/\sqrt{100} = (0.1 \ 0.1)'$. If we want to keep δ^* constant for $T = 200$ we have to set $\delta = \sqrt{T} \delta^* = \sqrt{200} (0.1 \ 0.1)' = (\sqrt{2} \ \sqrt{2})'$.

Without using δ^* explicitly, these computations are generalized for any two sample sizes by applying the normalization formula

$$\delta^{(2)} = \sqrt{T_2/T_1} \delta^{(1)}, \quad (4.5)$$

where T_i ($i = 1, 2$) are the sample sizes and $\delta^{(i)}$ ($i = 1, 2$) are the respective parameter vectors which assure that the level shifts have the same scaled shift magnitude in both sample sizes. In other words, $\delta^{(2)}$ is normalized with respect to $\delta^{(1)}$ and T_1 such that the same scaled magnitude is used for T_2 as for T_1 . Therefore, (4.5) is referred to as *normalization formula* and $\delta^{(2)}$ may be named as a normalized shift vector.

To assess the influence of the magnitude of the level shift I use the parameter values $\delta_i = 0$, $\delta_i = 1$, $\delta_i = 5$, $\delta_i = 10$, and $\delta_i = 20$ ($i = 1, \dots, n$) for both sample sizes $T = 100$ and $T = 200$. Of course, if $\delta_i = 0$ there occurs no shift and this case is regarded as a reference for evaluating the effects of a shift. Since the shift magnitudes are fixed for $T = 100$ and $T = 200$ the entries in the scaled parameter vector δ^* will decrease for $T = 200$ in line with the discussion of the scaling issue. In order to analyse the effects of increasing the sample size in case of level shifts that have the same scaled magnitude I also consider $\delta_i = \sqrt{2}$, $\delta_i = \sqrt{50}$, $\delta_i = \sqrt{200}$, and $\delta_i = \sqrt{800}$ ($i = 1, \dots, n$) for $T = 200$. These normalized quantities are obtained by applying formula (4.5) with $T_1 = 100$ and $T_2 = 200$. The vector $\delta^{(1)}$ contains the parameter values $\delta_i = 1$, $\delta_i = 5$, $\delta_i = 10$, and $\delta_i = 20$ ($i = 1, \dots, n$) respectively. Note that I will report the results with respect to $T = 200$ only in terms of some of the mentioned shift magnitudes.

To gauge the impact of the timing of the level shift three different relative break points are studied: $t_1 = 0.25$, $t_1 = 0.50$, and $t_1 = 0.75$ with $t_1 = T_1/T$. Relative shift dates very close to the beginning or the end of the sample are not considered in detail because the sample could be trimmed for the cointegration analysis by cutting off only a few of the first or last data points. Thus, one loses only a few observations. Nevertheless, I briefly comment on the effects of very early or very late shifts on the small sample size and power.

4.2 Size in Small Samples

In a first step the tests' sizes in small samples are analysed. I will use the term *size* for the frequency of rejecting the null hypothesis $H_0(r_0) : \text{rk}(\Pi) = r_0$ when the true cointegrating rank of the DGP is, in fact, r_0 . This use of the term *size* does not coincide with the exact definition of the size of a test. The definition would require to maximize the power function over the whole parameter space associated with the null hypothesis with respect to the nuisance parameters contained in the test statistics' small sample distributions. This is not done in my simulation framework.

As a point of departure I consider bivariate processes as described by (4.1) with $r = 1$, i.e. to evaluate the sizes of the tests the null hypothesis is $H_0(1) : \text{rk}(\Pi) = r_0 = 1$.

In Table 1 the results for $\psi_1 = 0.7$ and a shift in both components are shown for different shift magnitudes (δ) and values of the error term correlations (θ). As can be clearly seen the sizes of all tests approach zero if the magnitude of the shifts increases. Usually, a significant size reduction is observed for values equal to or larger than $\delta_1 = \delta_2 = 5$. Significant differences are indicated by bold type entries. Significance refers

Table 1. Rejection Frequencies of Tests for Bivariate DGPs with True Cointegrating Rank $r = 1$ ($\psi_1 = 0.7$), VAR Order $p = 1$, Sample Size $T = 100$, Nominal Significance Level 0.05, Rank under H_0 is $r_0 = 1$, Level Shift in all Components

| | $\delta = 0\nu_2$ | $\delta = 1\nu_2$ | $\delta = 5\nu_2$ | $\delta = 10\nu_2$ | $\delta = 20\nu_2$ | $\delta = 0\nu_2^1$ | $\delta = 1\nu_2$ | $\delta = 5\nu_2$ | $\delta = 10\nu_2$ | $\delta = 20\nu_2$ |
|---------------|-------------------|-------------------|-------------------|--------------------|--------------------|---------------------|-------------------|-------------------|--------------------|--------------------|
| | $\theta = 0$ | | | | | | $\theta = 0.8$ | | | |
| | $t_1 = 0.25$ | | | | | | | | | |
| LR_{Joh} | 0.0459 | 0.0450 | 0.0079 | 0.0000 | 0.0000 | 0.0622 | 0.0607 | 0.0389 | 0.0109 | 0.0030 |
| $LR_{S\&L}$ | 0.0474 | 0.0459 | 0.0266 | 0.0016 | 0.0000 | 0.0321 | 0.0315 | 0.0195 | 0.0047 | 0.0007 |
| LR_{Joh}^0 | 0.0475 | 0.0448 | 0.0132 | 0.0041 | 0.0008 | 0.0586 | 0.0591 | 0.0370 | 0.0167 | 0.0073 |
| $LR_{S\&L}^0$ | 0.0507 | 0.0549 | 0.0189 | 0.0006 | 0.0002 | 0.0537 | 0.0520 | 0.0418 | 0.0169 | 0.0067 |
| | $t_1 = 0.50$ | | | | | | | | | |
| LR_{Joh} | 0.0459 | 0.0440 | 0.0126 | 0.0006 | 0.0001 | 0.0622 | 0.0620 | 0.0426 | 0.0131 | 0.0028 |
| $LR_{S\&L}$ | 0.0474 | 0.0478 | 0.0385 | 0.0104 | 0.0011 | 0.0321 | 0.0318 | 0.0218 | 0.0098 | 0.0023 |
| LR_{Joh}^0 | 0.0475 | 0.0405 | 0.0038 | 0.0002 | 0.0000 | 0.0586 | 0.0572 | 0.0237 | 0.0060 | 0.0019 |
| $LR_{S\&L}^0$ | 0.0507 | 0.0529 | 0.0164 | 0.0015 | 0.0002 | 0.0537 | 0.0513 | 0.0393 | 0.0202 | 0.0095 |
| | $t_1 = 0.75$ | | | | | | | | | |
| LR_{Joh} | 0.0459 | 0.0444 | 0.0101 | 0.0001 | 0.0000 | 0.0622 | 0.0625 | 0.0374 | 0.0085 | 0.0019 |
| $LR_{S\&L}$ | 0.0474 | 0.0457 | 0.0275 | 0.0012 | 0.0000 | 0.0321 | 0.0315 | 0.0149 | 0.0036 | 0.0004 |
| LR_{Joh}^0 | 0.0475 | 0.0414 | 0.0025 | 0.0001 | 0.0000 | 0.0586 | 0.0560 | 0.0227 | 0.0039 | 0.0017 |
| $LR_{S\&L}^0$ | 0.0507 | 0.0534 | 0.0274 | 0.0048 | 0.0011 | 0.0537 | 0.0537 | 0.0495 | 0.0343 | 0.0236 |

Note: $\iota_2 = (1\ 1)'$. Bold entries indicate significant size differences with respect to the results for $\delta = (0\ 0)'$.

to differences that are equal to or larger than twice the standard error $s_{0.05} = 0.0022$ described above. The tests suggested by Saikkonen & Lütkepohl are slightly less affected by the size distortion in the sense that they lose less size than the Johansen tests given the same shift magnitude.

Comparing the different proposals concerning the deterministic terms it is seen that no test variant has an overall advantage over the other one. With respect to the various break points t_1 , significant differences for some δ are observed. Such significant differences also occur for the other DGPs considered in the following. But they do not exhibit a systematic pattern in the sense that specific break points generally cause a high or low size reduction. With respect to shift dates at the end of the sample I have observed similar results as for the other break points. For shifts at the very beginning of the sample, however, the size reduction is clearly less pronounced when the Johansen-type tests are considered. In fact, the size remains unchanged if the shift occurs, say, within the first 5% of the observations.

In a next step the simulations are repeated for DGPs where a shift only occurs in the nonstationary or stationary component respectively in order to compare the relative importance of these breaks. In Table 2 the results for a break in the stationary component display only some significant differences to the case of breaks in all time series of the DGP. The number of positive and negative deviations are approximately the same. Hence, the same conclusions as above apply. However, if a break occurs only in the nonstationary part all tests generally have a significantly larger size for shift magnitudes of $\delta_2 = 5$ or $\delta_2 = 10$ compared to the situation of a break in all components (compare Tables 1 and 3, bold type entries highlight significant differences). Accordingly, significant size reductions relative to the case of no shifts are rarely observed for shift magnitudes lower than $\delta_2 = 10$ or, in some cases, than $\delta_2 = 20$ when the tests suggested by Saikkonen & Lütkepohl are considered. Hence, a shift in the stationary part has a much more distortionary effect on the tests' size. This was also observed for higher dimensional processes. However, the size seems to approach zero for both types of level shifts. Still, one might argue that the rate at which the size reduces to zero is higher for shifts in the stationary part since the reduction is more pronounced given a certain shift magnitude. Interestingly, if a level shift is already present in the stationary part adding a shift to the nonstationary component does not cause further important size distortions.

Generalizing, one can conclude that level shifts which translate to shifts in the cointegration relations are much more important in terms of size distortion than nonorthogonal level shifts leading to outliers in a VECM only.

Table 2. Rejection Frequencies of Tests for Bivariate DGPs with True Cointegrating Rank $r = 1$ ($\psi_1 = 0.7$), VAR Order $p = 1$, Sample Size $T = 100$, Nominal Significance Level 0.05, Rank under H_0 is $r_0 = 1$, Level Shift Only in Stationary Component ($\delta_2 = 0$)

| | $\delta_1 = 0$ | $\delta_1 = 1$ | $\delta_1 = 5$ | $\delta_1 = 10$ | $\delta_1 = 20$ | $\delta_1 = 0$ | $\delta_1 = 1$ | $\delta_1 = 5$ | $\delta_1 = 10$ | $\delta_1 = 20$ |
|-----------------|----------------|----------------|----------------|-----------------|-----------------|----------------|----------------|----------------|-----------------|-----------------|
| | $\theta = 0$ | | | | | $\theta = 0.8$ | | | | |
| | $t_1 = 0.25$ | | | | | | | | | |
| $LR_{I_{oh}}$ | 0.0459 | 0.0445 | 0.0120 | 0.0005 | 0.0001 | 0.0622 | 0.0620 | 0.0395 | 0.0102 | 0.0026 |
| $LR_{S\&L}$ | 0.0474 | 0.0471 | 0.0235 | 0.0011 | 0.0000 | 0.0321 | 0.0301 | 0.0189 | 0.0036 | 0.0005 |
| $LR^0_{I_{oh}}$ | 0.0475 | 0.0470 | 0.0164 | 0.0039 | 0.0023 | 0.0586 | 0.0596 | 0.0484 | 0.0254 | 0.0128 |
| $LR^0_{S\&L}$ | 0.0507 | 0.0499 | 0.0226 | 0.0028 | 0.0010 | 0.0537 | 0.0544 | 0.0436 | 0.0135 | 0.0031 |
| | $t_1 = 0.50$ | | | | | | | | | |
| $LR_{I_{oh}}$ | 0.0459 | 0.0458 | 0.0184 | 0.0010 | 0.0000 | 0.0622 | 0.0610 | 0.0496 | 0.0116 | 0.0014 |
| $LR_{S\&L}$ | 0.0474 | 0.0476 | 0.0363 | 0.0091 | 0.0010 | 0.0321 | 0.0328 | 0.0244 | 0.0087 | 0.0019 |
| $LR^0_{I_{oh}}$ | 0.0475 | 0.0443 | 0.0046 | 0.0011 | 0.0004 | 0.0586 | 0.0585 | 0.0258 | 0.0096 | 0.0021 |
| $LR^0_{S\&L}$ | 0.0507 | 0.0513 | 0.0183 | 0.0032 | 0.0015 | 0.0537 | 0.0547 | 0.0387 | 0.0142 | 0.0049 |
| | $t_1 = 0.75$ | | | | | | | | | |
| $LR_{I_{oh}}$ | 0.0459 | 0.0449 | 0.0125 | 0.0005 | 0.0001 | 0.0622 | 0.0600 | 0.0419 | 0.0108 | 0.0019 |
| $LR_{S\&L}$ | 0.0474 | 0.0478 | 0.0252 | 0.0010 | 0.0000 | 0.0321 | 0.0332 | 0.0153 | 0.0027 | 0.0004 |
| $LR^0_{I_{oh}}$ | 0.0475 | 0.0431 | 0.0025 | 0.0003 | 0.0001 | 0.0586 | 0.0583 | 0.0241 | 0.0049 | 0.0003 |
| $LR^0_{S\&L}$ | 0.0507 | 0.0501 | 0.0264 | 0.0055 | 0.0040 | 0.0537 | 0.0537 | 0.0485 | 0.0275 | 0.0131 |

Note: Bold entries indicate significant differences with respect to corresponding entries in Table 1.

Table 3. Rejection Frequencies of Tests for Bivariate DGPs with True Cointegrating Rank $r = 1$, $\psi_1 = 0.7$, VAR Order $p = 1$, Sample Size $T = 100$, Nominal Significance Level 0.05, Rank under H_0 is $r_0 = 1$, Level Shift Only in Nonstationary Component ($\delta_1 = 0$)

| | $\delta_2 = 0$ | $\delta_2 = 1$ | $\delta_2 = 5$ | $\delta_2 = 10$ | $\delta_2 = 20$ | $\delta_2 = 0$ | $\delta_2 = 1$ | $\delta_2 = 5$ | $\delta_2 = 10$ | $\delta_2 = 20$ |
|---------------|----------------|----------------|----------------|-----------------|-----------------|----------------|----------------|----------------|-----------------|-----------------|
| | $\theta = 0$ | | | | | $\theta = 0.8$ | | | | |
| | $t_1 = 0.25$ | | | | | | | | | |
| LR_{Joh} | 0.0459 | 0.0474 | 0.0357 | 0.0159 | 0.0005 | 0.0622 | 0.0616 | 0.0445 | 0.0152 | 0.0003 |
| $LR_{S\&L}$ | 0.0474 | 0.0466 | 0.0482 | 0.0456 | 0.0124 | 0.0321 | 0.0327 | 0.0374 | 0.0356 | 0.0094 |
| LR_{Joh}^0 | 0.0475 | 0.0459 | 0.0392 | 0.0204 | 0.0002 | 0.0586 | 0.0578 | 0.0456 | 0.0222 | 0.0010 |
| $LR_{S\&L}^0$ | 0.0507 | 0.0552 | 0.0459 | 0.0351 | 0.0048 | 0.0537 | 0.0477 | 0.0482 | 0.0364 | 0.0036 |
| | $t_1 = 0.50$ | | | | | | | | | |
| LR_{Joh} | 0.0459 | 0.0452 | 0.0347 | 0.0168 | 0.0008 | 0.0622 | 0.0615 | 0.0443 | 0.0145 | 0.0004 |
| $LR_{S\&L}$ | 0.0474 | 0.0475 | 0.0467 | 0.0418 | 0.0126 | 0.0321 | 0.0316 | 0.0360 | 0.0326 | 0.0120 |
| LR_{Joh}^0 | 0.0475 | 0.0441 | 0.0318 | 0.0114 | 0.0004 | 0.0586 | 0.0571 | 0.0421 | 0.0127 | 0.0004 |
| $LR_{S\&L}^0$ | 0.0507 | 0.0548 | 0.0486 | 0.0331 | 0.0078 | 0.0537 | 0.0500 | 0.0486 | 0.0370 | 0.0074 |
| | $t_1 = 0.75$ | | | | | | | | | |
| LR_{Joh} | 0.0459 | 0.0453 | 0.0378 | 0.0170 | 0.0009 | 0.0622 | 0.0607 | 0.0469 | 0.0161 | 0.0002 |
| $LR_{S\&L}$ | 0.0474 | 0.0480 | 0.0487 | 0.0412 | 0.0120 | 0.0321 | 0.0319 | 0.0364 | 0.0350 | 0.0083 |
| LR_{Joh}^0 | 0.0475 | 0.0462 | 0.0307 | 0.0111 | 0.0002 | 0.0586 | 0.0584 | 0.0366 | 0.0130 | 0.0006 |
| $LR_{S\&L}^0$ | 0.0507 | 0.0531 | 0.0530 | 0.0397 | 0.0085 | 0.0537 | 0.0532 | 0.0506 | 0.0395 | 0.0088 |

Note: Bold entries indicate significant differences with respect to corresponding entries in Table 1.

Table 4. Rejection Frequencies of Tests for Bivariate DGPs with True Cointegrating Rank $r = 1$ ($\psi_1 = 0.7$), $\theta = 0.8$, VAR Order $p = 1$, Sample Size $T = 200$, Nominal Significance Level 0.05, $H_0(1) : r_0 = 1$, Level Shift in all Components

| | $\delta = 0\iota_2$ | $\delta = \sqrt{2}\iota_2$ | $\delta = 5\iota_2$ | $\delta = \sqrt{50}\iota_2$ | $\delta = 10\iota_2$ | $\delta = \sqrt{200}\iota_2$ | $\delta = \sqrt{800}\iota_2$ |
|---------------|---------------------|----------------------------|---------------------|-----------------------------|----------------------|------------------------------|------------------------------|
| | $t_1 = 0.25$ | | | | | | |
| LR_{Joh} | 0.0537 | 0.0528 | 0.0467 | 0.0393 | 0.0262 | 0.0127 | 0.0027 |
| $LR_{S\&L}$ | 0.0365 | 0.0348 | 0.0217 | 0.0163 | 0.0086 | 0.0031 | 0.0007 |
| LR_{Joh}^0 | 0.0538 | 0.0534 | 0.0470 | 0.0375 | 0.0240 | 0.0129 | 0.0055 |
| $LR_{S\&L}^0$ | 0.0539 | 0.0522 | 0.0487 | 0.0445 | 0.0285 | 0.0160 | 0.0052 |
| | $t_1 = 0.50$ | | | | | | |
| LR_{Joh} | 0.0537 | 0.0538 | 0.0484 | 0.0417 | 0.0298 | 0.0150 | 0.0022 |
| $LR_{S\&L}$ | 0.0365 | 0.0356 | 0.0256 | 0.0185 | 0.0109 | 0.0063 | 0.0008 |
| LR_{Joh}^0 | 0.0538 | 0.0516 | 0.0385 | 0.0232 | 0.0114 | 0.0050 | 0.0016 |
| $LR_{S\&L}^0$ | 0.0539 | 0.0542 | 0.0504 | 0.0436 | 0.0274 | 0.0168 | 0.0085 |
| | $t_1 = 0.75$ | | | | | | |
| LR_{Joh} | 0.0537 | 0.0514 | 0.0423 | 0.0345 | 0.0215 | 0.0093 | 0.0018 |
| $LR_{S\&L}$ | 0.0365 | 0.0332 | 0.0190 | 0.0102 | 0.0039 | 0.0013 | 0.0002 |
| LR_{Joh}^0 | 0.0538 | 0.0532 | 0.0384 | 0.0235 | 0.0107 | 0.0050 | 0.0012 |
| $LR_{S\&L}^0$ | 0.0539 | 0.0532 | 0.0512 | 0.0480 | 0.0420 | 0.0344 | 0.0175 |

Note: $\iota_2 = (1\ 1)'$. Bold entries indicate significant size differences with respect to the results for $\delta = (0\ 0)'$.

The impact of the parameters θ and ψ_1 on the empirical size level is in general more important than their effect on the way the size is distorted. For increasing δ , one may observe that high correlation leads to a slower size reduction in case of $LR_{S\&L}$ and $LR_{S\&L}^0$ and to a faster size reduction in case of LR_{Joh} and LR_{Joh}^0 . However, this may be an outcome of the different size levels caused by θ . Setting $\psi_1 = 0.9$ produces lower empirical sizes for all tests but the general conclusions regarding the size distortion remain largely the same. Therefore, these results are not presented here.

Some further issues that relate to the dependency of the outcomes on the sample size, the dimension of the DGP, and the number of time series affected by a level shift are discussed in the following.

In Table 4 results for the case of $T = 200$ with a shift in all components are collected. The parameters of the DGPs are the same as in Table 1 with the exception of the shift

magnitudes.⁵ In Table 4 the corresponding magnitudes according to the normalization formula (4.5) are used such that the scaled shift magnitude is the same for $T = 100$ and $T = 200$. Additionally, the results for $\delta = (5 \ 5)'$ and $\delta = (10 \ 10)'$ are included in order to enable a comparison with $T = 100$ for the case of a fixed shift magnitude. Mainly, the tests' size significantly fall for $\delta = (5 \ 5)'$ and higher values of δ . Again, the Saikkonen-Lütkepohl proposals are slightly less affected by the size distortion. The magnitude of 5 corresponds to $\delta = (3.5 \ 3.5)'$ for $T = 100$ in line with (4.5). Above I have described that the sizes fall significantly for $\delta = (5 \ 5)'$ or higher values if $T = 100$. This would mean that a significant size reduction occurs for lower normalized shift magnitudes if the sample size increases. But comparing the results for $\delta = (5 \ 5)'$ we can clearly see that the tests' sizes are higher and the size reduction is lower if $T = 200$. We may also conclude that a significant reduction would have occurred for shift magnitudes lower than five if $T = 100$ is considered. Hence, when taking into account the scaling issue the relative outcomes are the same in case of the larger sample size. On the other hand, comparing the findings for $\delta = (5 \ 5)'$ and $\delta = (10 \ 10)'$ it follows that a shift of a given magnitude may distort the size less for increasing sample sizes. Nevertheless, if the magnitude of the shift increases the empirical sizes still approach zero.

I have also performed simulation experiments for four-dimensional DGPs to check the effects of ignoring level shifts in higher-dimensional and more complicated processes. As an example the results for a DGP with a cointegrating rank $r = 2$ and a break in all components are shown in Table 5. It can be seen that a significant size reduction occurs for $\delta = (5 \ 5 \ 5 \ 5)'$ and that the size decreases to zero for higher values of δ . Note that in two cases a significant fall of the size of LR_{Joh}^0 can already be observed for $\delta = (1 \ 1 \ 1 \ 1)'$. Nevertheless, the differences concerning the various test proposals, the deterministic terms and the relative break points are less marked than in the bivariate setup. Interestingly, this can also be observed for other four-dimensional processes not shown here. However, as in the bivariate setup a level shift in the stationary components is much more important than a shift in the nonstationary part. If breaks are already present in the stationary time series then adding a level shift to the nonstationary components does not exacerbate the tests' size properties. Since the findings are similar for both bivariate and four-dimensional DGPs detailed results are not given here.

The last remark already refers to the issue whether it is important how many components are affected by a shift. A first conclusion is that the type of the shift is of interest in this respect, i.e. whether the break occurs in the nonstationary or stationary part. I can

⁵In Table 4 only the results for $\theta = 0.8$ are shown.

Table 5. Rejection Frequencies of Tests for Four-dimensional DGPs with True Cointegrating Rank $r = 2$ ($\psi_1 = \psi_2 = 0.7$), VAR Order $p = 1$, Sample Size $T = 100$, Nominal Significance Level 0.05, Rank under H_0 is $r_0 = 2$, Level Shift in all Components

| | $\delta = 0\iota_4$ | $\delta = 1\iota_4$ | $\delta = 2\iota_4$ | $\delta = 5\iota_4$ | $\delta = 10\iota_4$ | $\delta = 0\iota_4$ | $\delta = 1\iota_4$ | $\delta = 2\iota_4$ | $\delta = 5\iota_4$ | $\delta = 10\iota_4$ |
|---------------|--|---------------------|---------------------|---------------------|----------------------|--|---------------------|---------------------|---------------------|----------------------|
| | $\Theta = [(0\,0)^\vee : (0\,0)^\vee]$ | | | | | $\Theta = [(0.4\,0.4)^\vee : (0.4\,0.4)^\vee]$ | | | | |
| | $t_1 = 0.25$ | | | | | | | | | |
| LR_{Joh} | 0.0239 | 0.0226 | 0.0043 | 0.0019 | 0.0011 | 0.0473 | 0.0438 | 0.0189 | 0.0063 | 0.0031 |
| $LR_{S\&L}$ | 0.0310 | 0.0300 | 0.0073 | 0.0050 | 0.0034 | 0.0360 | 0.0331 | 0.0151 | 0.0046 | 0.0030 |
| LR_{Joh}^0 | 0.0319 | 0.0281 | 0.0092 | 0.0054 | 0.0038 | 0.0494 | 0.0484 | 0.0221 | 0.0112 | 0.0077 |
| $LR_{S\&L}^0$ | 0.0463 | 0.0443 | 0.0124 | 0.0062 | 0.0052 | 0.0532 | 0.0562 | 0.0278 | 0.0141 | 0.0068 |
| | $t_1 = 0.50$ | | | | | | | | | |
| LR_{Joh} | 0.0239 | 0.0216 | 0.0033 | 0.0009 | 0.0006 | 0.0473 | 0.0453 | 0.0184 | 0.0048 | 0.0011 |
| $LR_{S\&L}$ | 0.0310 | 0.0292 | 0.0138 | 0.0068 | 0.0067 | 0.0360 | 0.0331 | 0.0186 | 0.0072 | 0.0051 |
| LR_{Joh}^0 | 0.0319 | 0.0271 | 0.0048 | 0.0025 | 0.0020 | 0.0494 | 0.0471 | 0.0174 | 0.0074 | 0.0035 |
| $LR_{S\&L}^0$ | 0.0463 | 0.0422 | 0.0154 | 0.0092 | 0.0081 | 0.0532 | 0.0546 | 0.0337 | 0.0180 | 0.0106 |
| | $t_1 = 0.75$ | | | | | | | | | |
| LR_{Joh} | 0.0239 | 0.0204 | 0.0032 | 0.0014 | 0.0011 | 0.0473 | 0.0454 | 0.0165 | 0.0057 | 0.0024 |
| $LR_{S\&L}$ | 0.0310 | 0.0299 | 0.0094 | 0.0056 | 0.0043 | 0.0360 | 0.0329 | 0.0141 | 0.0056 | 0.0016 |
| LR_{Joh}^0 | 0.0319 | 0.0246 | 0.0015 | 0.0005 | 0.0004 | 0.0494 | 0.0481 | 0.0125 | 0.0040 | 0.0023 |
| $LR_{S\&L}^0$ | 0.0463 | 0.0443 | 0.0138 | 0.0075 | 0.0059 | 0.0532 | 0.0536 | 0.0316 | 0.0170 | 0.0117 |

Note: $\iota_4 = (1\ 1\ 1\ 1)'$. Bold entries indicate significant size differences with respect to the results for $\delta = (0\ 0\ 0\ 0)'$.

Table 6. Rejection Frequencies of Tests for Four-dimensional DGPs with Cointegrating Rank $r = 2$ ($\psi_1 = \psi_2 = 0.7$), $\Theta = [(0.4 \ 0.4)' : (0.4 \ 0.4)']$, VAR Order $p = 1$, Sample Size $T = 100$, $t_1 = 0.50$, Nominal Significance Level 0.05, Level Shift in both Nonstationary Components

| | $\delta_3 = \delta_4 = 0$ | $\delta_3 = \delta_4 = 1$ | $\delta_3 = \delta_4 = 5$ | $\delta_3 = \delta_4 = 10$ | $\delta_3 = \delta_4 = 20$ |
|---------------|---------------------------|---------------------------|---------------------------|----------------------------|----------------------------|
| LR_{Joh} | 0.0473 | 0.0453 | 0.0281 | 0.0119 | 0.0050 |
| $LR_{S\&L}$ | 0.0360 | 0.0353 | 0.0378 | 0.0310 | 0.0253 |
| LR_{Joh}^0 | 0.0494 | 0.0476 | 0.0342 | 0.0177 | 0.0086 |
| $LR_{S\&L}^0$ | 0.0532 | 0.0533 | 0.0521 | 0.0435 | 0.0300 |

Note: Bold entries indicate significant size differences with respect to the results for $\delta = (0 \ 0 \ 0 \ 0)'$.

conclude from simulation experience of which results are not presented here that it is of less importance how many stationary components are affected by a shift. If one stationary component includes a level shift introducing a shift to a previously unaffected stationary component does not change the tests' size properties significantly. In contrast, increasing the number of affected nonstationary time series reduces the tests' size significantly in some cases.

Among the previous studies mentioned in Section 3.1 only Doornik et al. (1998) analyse the small sample size properties in case of neglected breaks. Taking into account that the level shifts in the nonstationary part are less important and interpreting the unrestricted impulse dummies used by Doornik et al. (1998) as level shifts orthogonal to the cointegration space it is not surprising that they report reasonable size properties for breaks of a smaller magnitude. However, with respect to their large dummy variables with a magnitude of fifty and thirty times the standard deviation of the error terms we may expect a pronounced reduction in the tests' sizes since these magnitudes are much larger than the ones analysed in my study. In contrast, Doornik et al. (1998) found only a slight size decrease.

I also observe rather low empirical size values, at least for the Johansen tests, when examining a four-dimensional DGP comparable to the ones employed in Doornik et al. (1998). The DGP has a true cointegrating rank of $r = 2$, strong error term correlation ($\Theta = [(0.4 \ 0.4)' : (0.4 \ 0.4)']$), $t_1 = 0.50$ and the level shifts occurs in the two nonstationary components, i.e. I consider nonorthogonal level shifts. The low size values displayed in Table 6 are an outcome of clearly significant size reductions compared to $\delta = (0 \ 0 \ 0 \ 0)'$. Therefore, one may conclude that the effects of ignoring shifts seem to be less important

within the framework used in Doornik et al. (1998). This conclusion may point again to the fact that unrestricted impulse dummies considered by Doornik et al. (1998) could just be the result of simple one-time outliers in levels instead of being the outcome of nonorthogonal level shifts. For one-time outliers we may expect a much lower distortionary effect on the tests' small sample sizes.

4.3 Small Sample Power

In the second part of the Monte Carlo study I examine the small sample power of the cointegration tests. The small sample power is not size adjusted since such an adjustment is usually not possible in applied work. Again, I start by analysing bivariate processes with shifts in both components. The true cointegrating rank is $r = 1$ ($|\psi_1| < 1$) and the null hypothesis is $H_0(0) : r_0 = 0$. In Figure 1 the small sample power is depicted depending on $\delta = (\delta_1 \ \delta_2)$ with $\delta_1 = \delta_2$ for $T = 100$ and different values of ψ_1 and θ . I only show the small sample power for the break point $t_1 = 0.50$ since the power differences caused by the various break dates are generally of minor importance. The same result was also found by Inoue (1999) for the Johansen-type tests used in his analysis. However, in line with the results for the size the small sample power of the Johansen tests is not reduced for shifts occurring at the very beginning of the sample whereas shifts at the end of the sample produce similar results as the break point $t_1 = 0.50$.

For $\delta = (1 \ 1)'$ the power is almost identical to the situation of no level shifts ($\delta = (0 \ 0)'$). Increasing the magnitude of the shifts to five reduces the power of the tests clearly. The tests lose sometimes more than 50 percent of their original power and the power decreases further if δ_i ($i = 1, 2$) is set to 10 or 20. However, there are differences in the reduction among the tests indicated by the steepness of the power curves. For example, $LR_{S\&L}$ loses relatively less power compared to all other tests, especially in the absence of innovation correlation. It also turns out that the tests allowing for a trend loses less power compared to the test versions assuming $\mu_1 = 0$. As a result of this advantage the tests accommodating a linear trend have a higher power for $\delta_i \geq 5$ ($i = 1, 2$) in contrast to the situations of $\delta = (0 \ 0)'$ or $\delta = (1 \ 1)'$. This was also observed by Gregory & Hansen (1996) for single-equation cointegration tests with and without a linear trend.⁶ The smaller power loss could be due to the trend term that may be able to capture some effects of the level shift.

By and large, the parameters ψ_1 and θ just affect the level of the power. Nevertheless, for $\theta = 0$ a higher value of ψ_1 causes a steeper power curve between $\delta = (1 \ 1)'$ and

⁶As already mentioned, Doornik et al. (1998) analyse only the Johansen test with a trend and Inoue (1999) applies only one Johansen-type test to the level shift case.

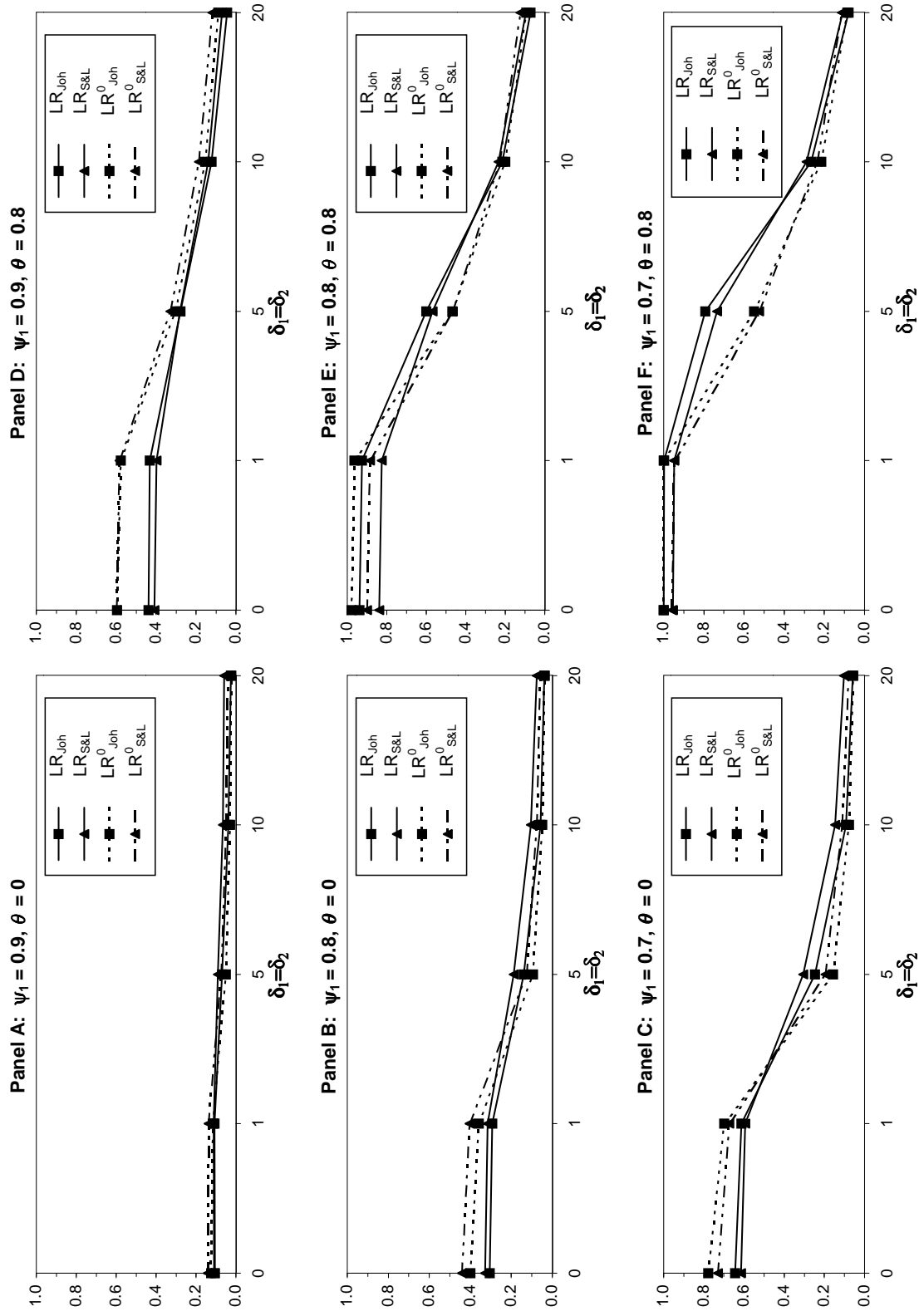


Figure 1. Small sample power of tests for $H_0(0) : r_0 = 0$ based on bivariate DGPs with true cointegrating rank $r = 1$, $T = 100$, $t_1 = 0.5$, level shift in all components.

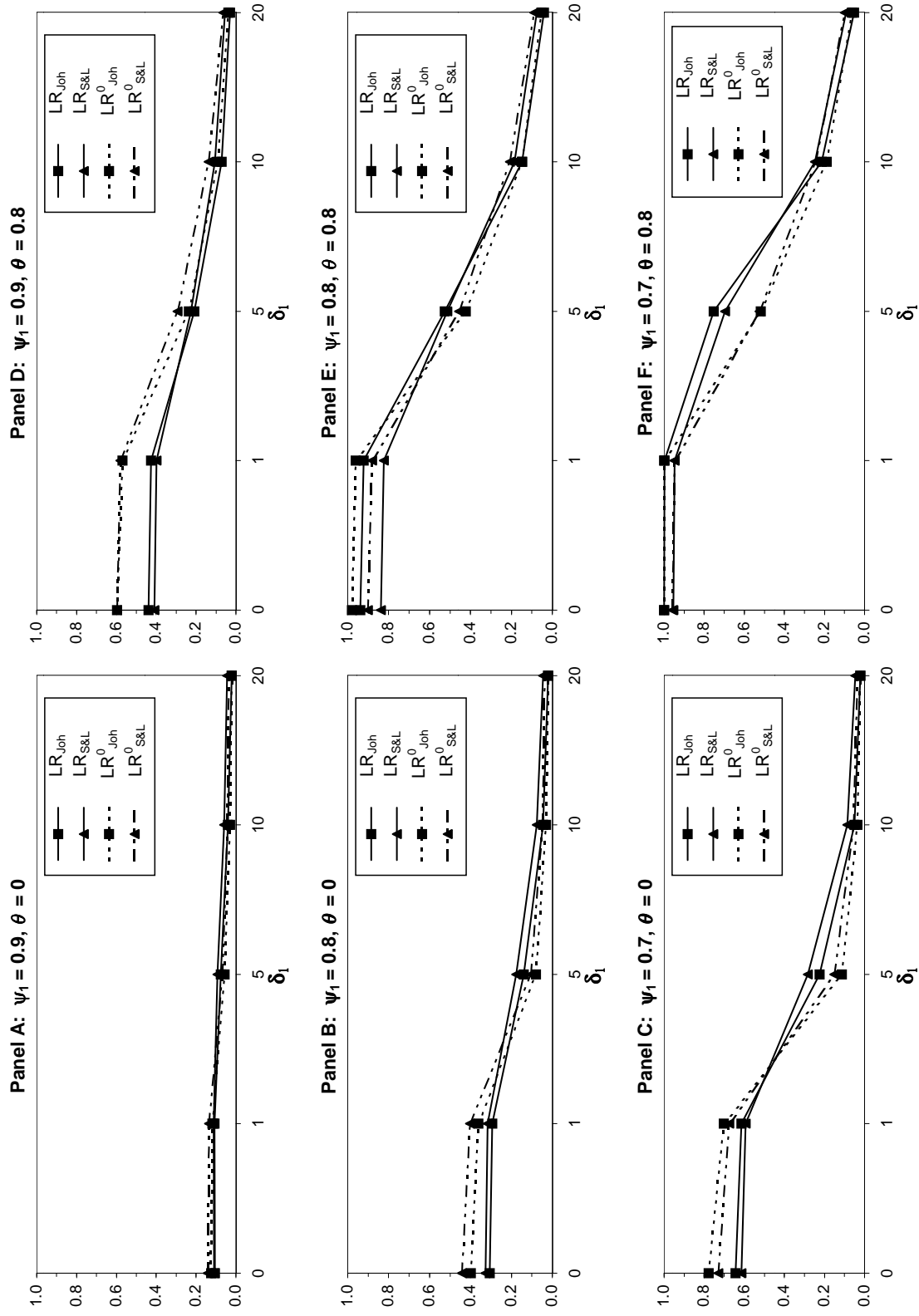


Figure 2. Small sample power of tests for $H_0(0) : r_0 = 0$ based on bivariate DGPs with true cointegrating rank $r = 1$, $T = 100$, $t_1 = 0.5$, level shift only in stationary component.

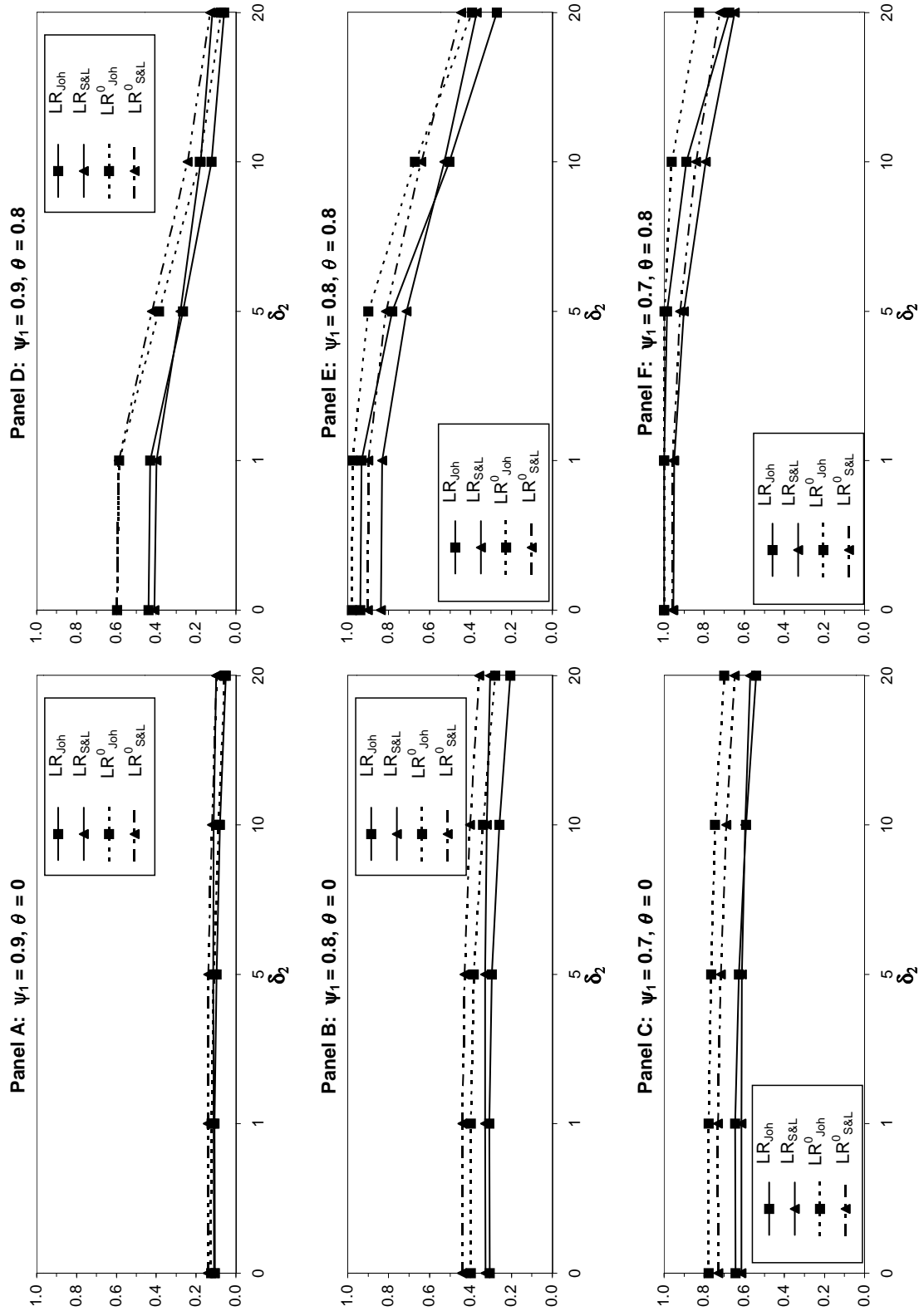


Figure 3. Small sample power of tests for $H_0(0) : r_0 = 0$ based on bivariate DGPs with true cointegrating rank $r = 1$, $T = 100$, $t_1 = 0.5$, level shift only in nonstationary component.

$\delta = (5 \ 5)'$. But independent of ψ_1 the power of all tests seem to converge to the same power level for large values of δ_i ($i = 1, 2$). If $\delta = (20 \ 20)'$, the power varies around 5% ($\theta = 0$) and 10% ($\theta = 0.8$).

As for the tests' sizes, the loss in small sample power depends on the component in which the shift occurs. In Figures 2 and 3 the power is shown for a shift only occurring in the stationary or the nonstationary component respectively. If the nonstationary component is affected by a level shift in case of no error term correlation the power remains almost unchanged independent of the shift magnitude. And even if the nonstationary component is related to the stationary part via the error term correlation the loss in power is less marked compared to the situation of a shift in the stationary component. Furthermore, the power curves for the case of a break in the stationary component only are almost identical to the those referring to shifts in both components. That means, if the stationary component is affected by a shift, adding a shift to the nonstationary part of the process does not change the power properties of the cointegration tests.

Also for a sample size of $T = 200$ shifts in both components lead to a dramatic fall in the power (see Figure 4). Although the tests have now a higher power if $\delta = (0 \ 0)'$, the small sample power reduces to approximately the same values as for $T = 100$ when using the corresponding higher values of δ obtained from the normalization (4.5). Comparing the rejection frequencies for $\delta = (5 \ 5)'$ and $\delta = (10 \ 10)'$, it is seen that for a given shift magnitude the relative power loss is often less important if the sample size increases. As in the case of $T = 100$ the test versions allowing for a linear trend are less affected by the power reduction. Their power advantage is particularly pronounced for $\psi_1 = 0.8$ or $\psi_1 = 0.7$.

I have also simulated four-dimensional DGPs with $r = 1$ and $r = 2$ for $T = 100$ with a shift in all time series. For the DGPs with $r = 1$ one gets similar results as in the bivariate setup. Therefore, the corresponding graph is not shown. However, testing the null hypothesis $r_0 = 0$ if $r = 2$ the tests still have a small sample power between 20% and 40% for $\delta = (20, 20, 20, 20)$ (see Figure 5). Hence, there is at least some chance of finding a cointegration rank of one if there are two cointegration relations present. On the other hand, finding a second cointegration relation, i.e. rejecting $H(0) : r_0 = 1$, is very unlikely as shown in Figure 6. In contrast to the bivariate DGPs the differences in the small sample power with respect to the assumptions made about the deterministic terms are less pronounced within the higher dimensional framework. The same is true regarding the differences in the performance of the test proposals by Johansen and Saikkonen & Lütkepohl.

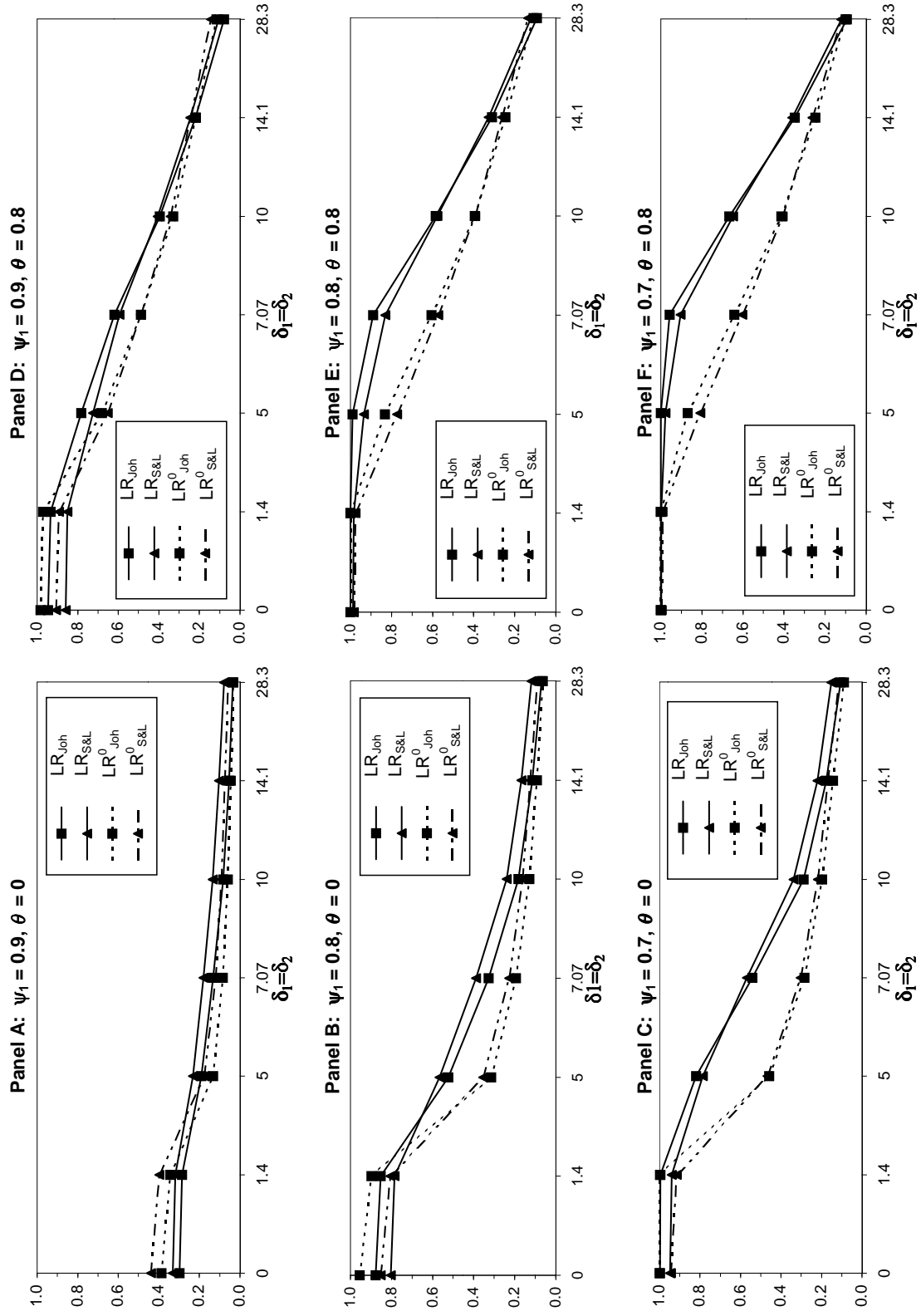


Figure 4. Small sample power of tests for $H_0(0) : r_0 = 0$ based on bivariate DGPs with true cointegrating rank $r = 1$, $T = 200$, $t_1 = 0.5$, level shift in all components.

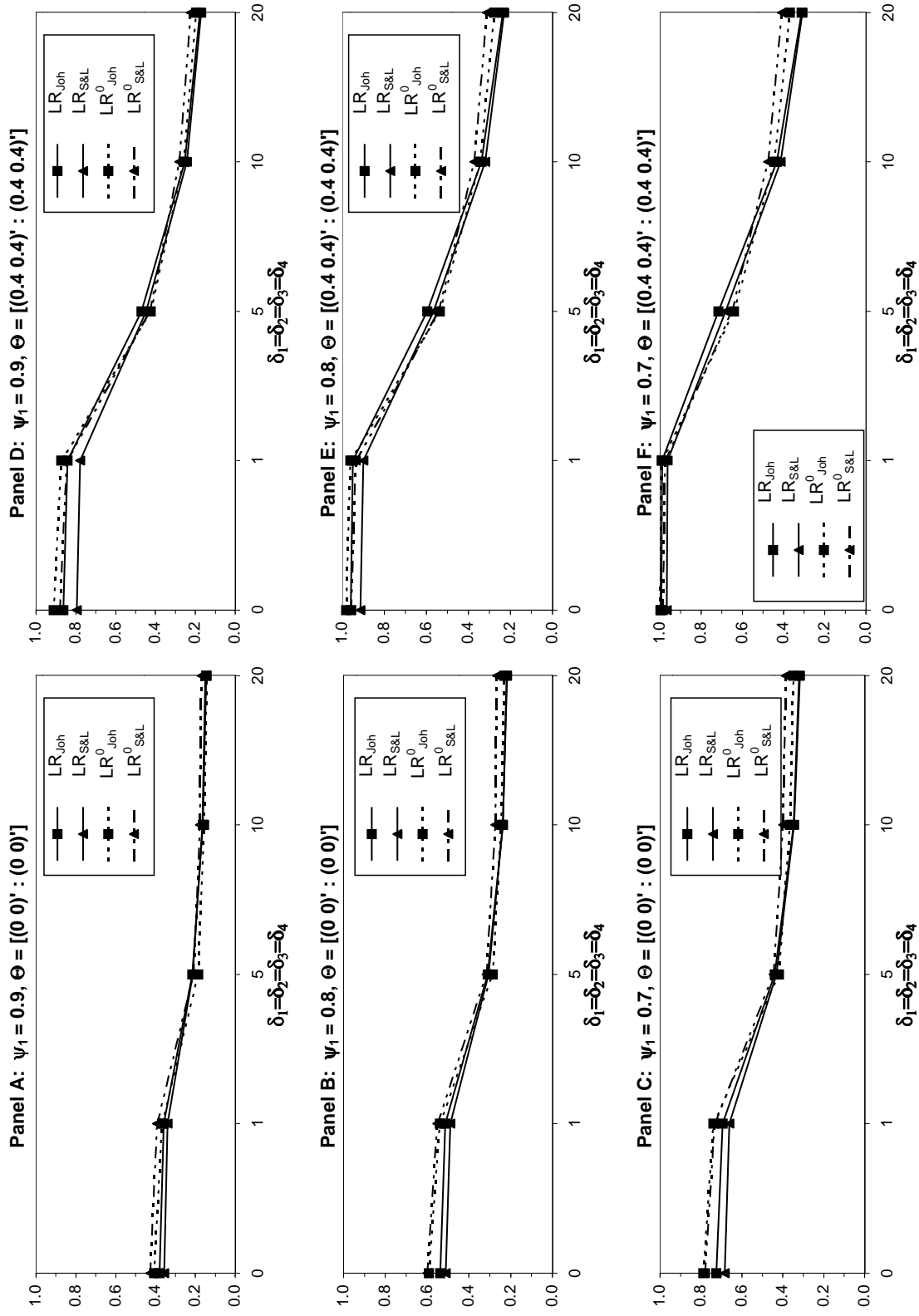


Figure 5. Small sample power of tests for $H_0(0) : r_0 = 0$ based on four-dimensional DGPs with true cointegrating rank $r = 2$ ($\psi_2 = 0.7$), $T = 100$, $t_1 = 0.5$, level shift in all components.

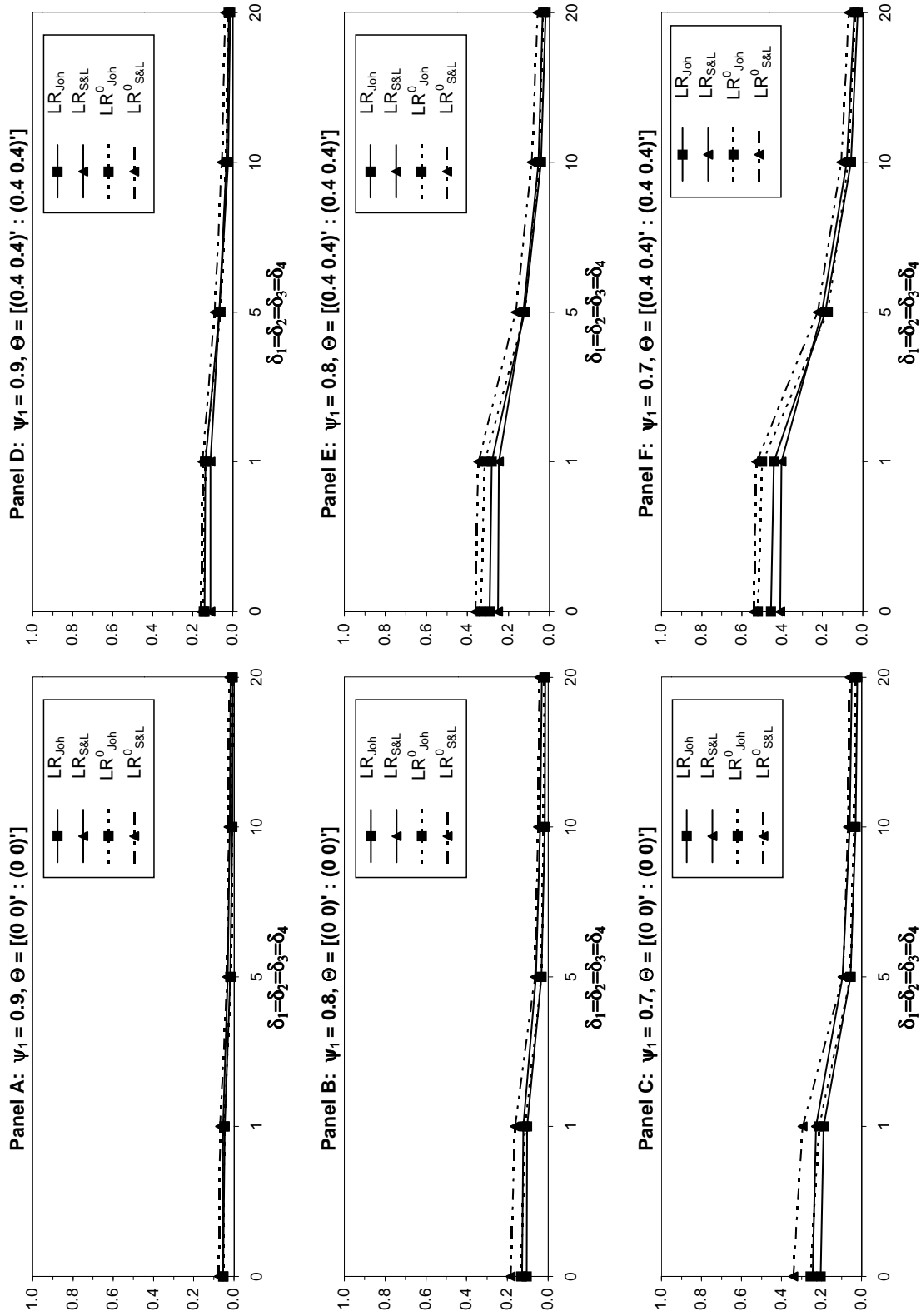


Figure 6. Small sample power of tests for $H_0(1) : r_0 = 1$ based on four-dimensional DGPs with true cointegrating rank $r = 2$ ($\psi_2 = 0.7$), $T = 100$, $t_1 = 0.5$ level shift in all components.

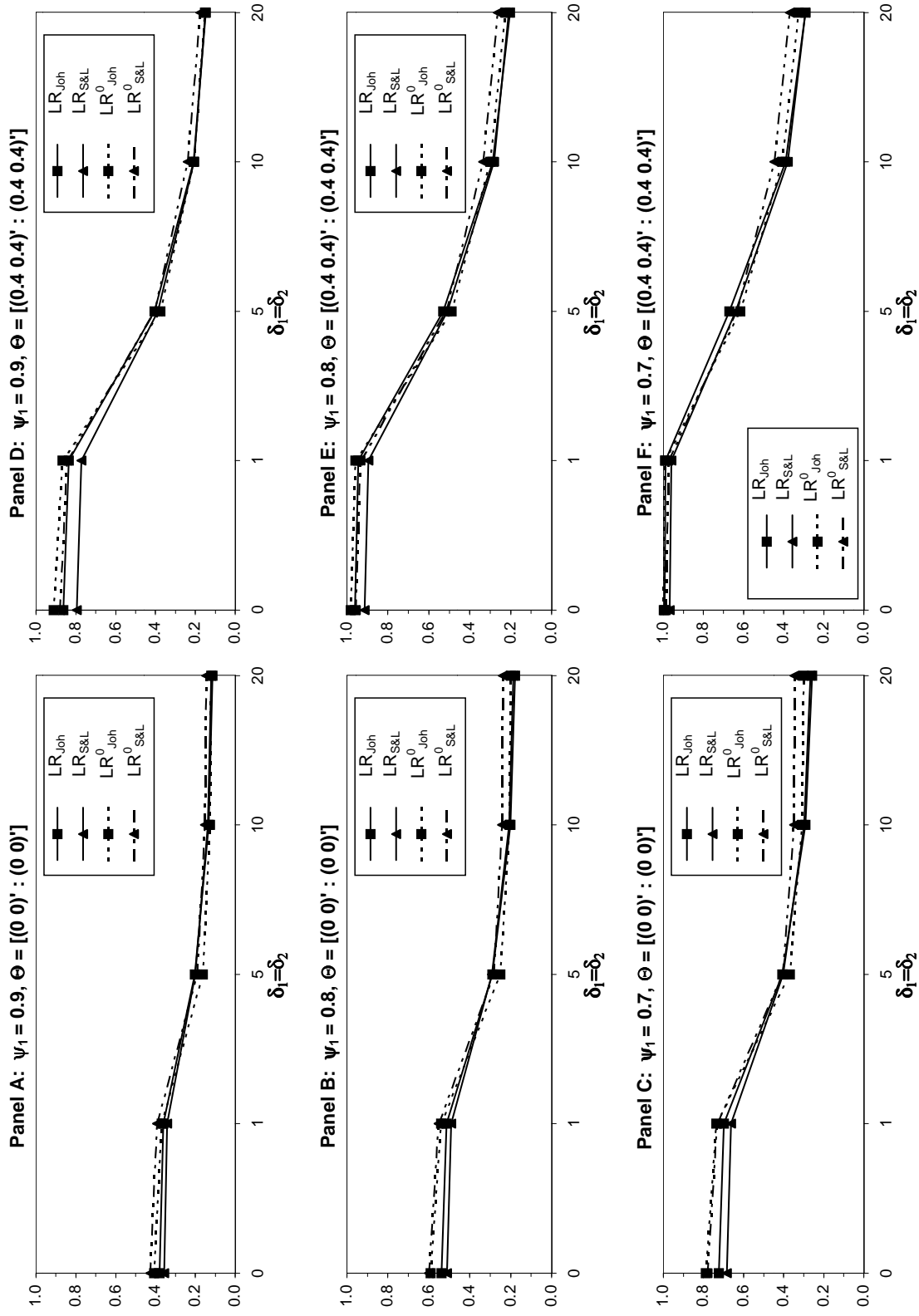


Figure 7. Small sample power of tests for $H_0(0) : r_0 = 0$ based on four-dimensional DGPs with true cointegrating rank $r = 2$ ($\psi_2 = 0.7$), $T = 100$, $t_1 = 0.5$, level shift in both stationary components.

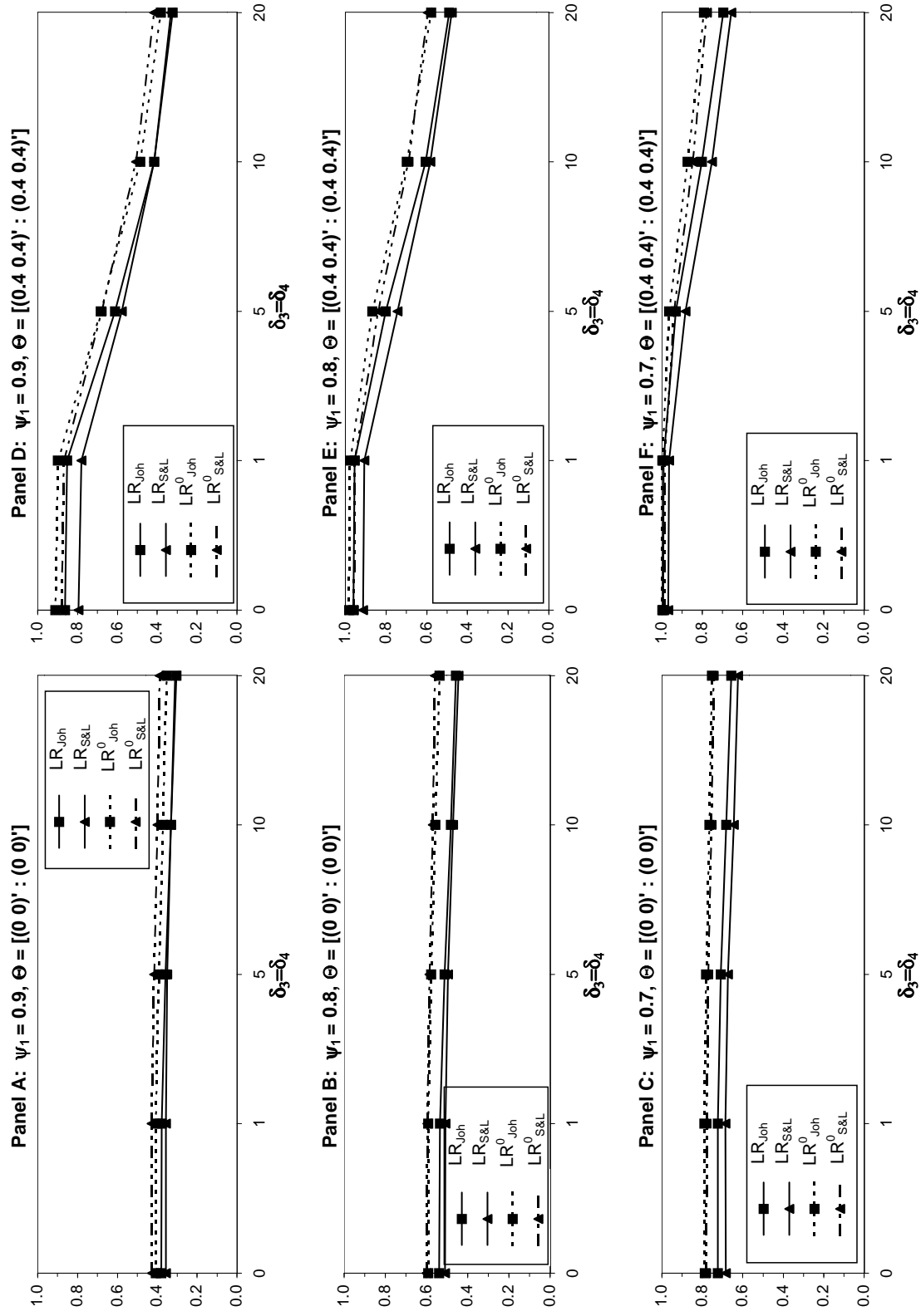


Figure 8. Small sample power of tests for $H_0(0) : r_0 = 0$ based on four-dimensional DGPs with true cointegrating rank $r = 2$ ($\psi_2 = 0.7$), $T = 100$, $t_1 = 0.5$, level shift in both nonstationary components.

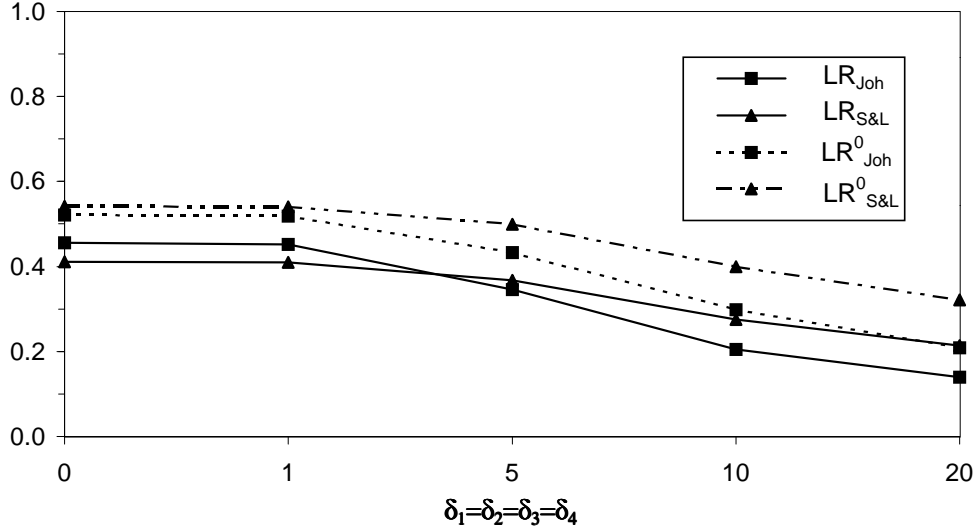


Figure 9. Small sample power of tests for $H_0(0) : r_0 = 1$ based on four-dimensional DGPs with true cointegrating rank $r = 2$ ($\psi_1 = \psi_2 = 0.7$), $\Theta = [(0.4 \ 0.4)' : (0.4 \ 0.4)']$, $T = 100$, $t_1 = 0.5$, level shift in both nonstationary components.

Focusing on the DGPs with $r = 2$ I have analysed again the relative importance of shifts occurring either in the stationary or nonstationary components. As in the bivariate setup, shifts in the stationary part of the DGP that translate to shifts in the cointegration relations are much more important in terms of power reduction than breaks in the nonstationary part that are orthogonal to the cointegration space (compare Figures 7 and 8). In fact, the relative rejection frequencies in case of shifts in the two stationary components and in case of breaks in all time series are almost the same. In contrast, shifts in both nonstationary components lead to clearly less power reduction. These findings regarding the relative importance of shifts in the stationary and nonstationary components may explain why Gregory & Hansen (1996) and Inoue (1999) report a clear loss in the small sample power of the single-equation and Johansen-type tests they consider in case of ignoring level shifts in the cointegration relations whereas Doornik et al. (1998) only observe smaller power losses for large unrestricted impulse dummies compared to the situation of small dummies. Again, it is assumed in this respect that a level shift in a cointegration relation corresponds to a level shift in the time series which is orthogonal to the cointegration space, i.e. to a shift in a stationary component of the DGP I used. Accordingly, an unrestricted impulse dummy is regarded to be equivalent to a nonorthogonal shift, i.e. to a shift in a nonstationary component.

I observe, however, that the power loss reported in Doornik et al. (1998) is even lower than for a comparable DGP with orthogonal level shifts in accordance with my framework.

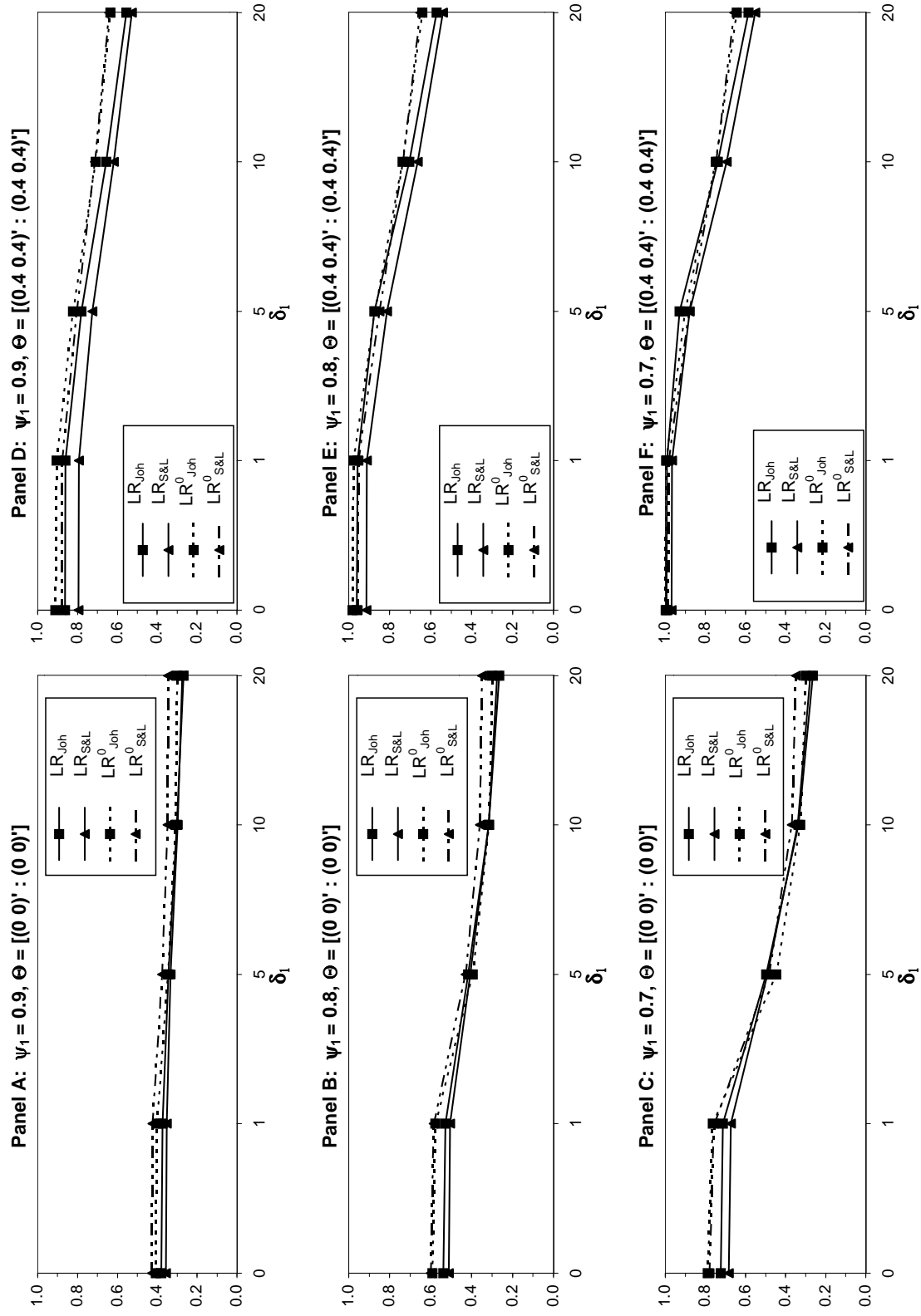


Figure 10. Small sample power of tests for $H_0(0) : r_0 = 0$ based on four-dimensional DGPs with true cointegrating rank $r = 2$ ($\psi_2 = 0.7$), $T = 100$, $t_1 = 0.5$, level shift in the first stationary component.

Analysing a four-dimensional DGP with $r = 2$ and an unrestricted impulse dummy in two components and testing $H_0(r_0) = 1$ Doornik et al. (1998) find that the LR_{Joh} test loses only one third of its original small sample power when the magnitude of the unrestricted impulse dummies is increased from three and five times to thirty and fifty times the standard deviation of the error terms respectively. For the comparable DGP I consider it can be seen in Figure 9 that the Johansen tests lose about two third, the Saikkonen-Lütkepohl tests about one half of their original power when a shift magnitude of twenty times the standard deviation is considered in comparison to the case of no level shifts. But note again, that unrestricted impulse dummies could also be the effect of one-time outliers. In the same way as for the tests' small sample size we may expect that such outliers are less important and this may explain the differences in the observed outcomes.

Using the four-dimensional setup with $r = 2$ I also examine whether it is important how many time series are affected by a level shift. In contrast to the outcome for the small sample size, it is crucial for the power how many stationary components are affected by a shift. Comparing Figures 7 and 10 an obvious power decline can be observed when a shift is additionally introduced to the second stationary component. On the other hand, whether one or two nonstationary components are hit by a level shift is of minor relevance. Therefore, the respective graphs are not shown here. Furthermore, if there already exist breaks in the stationary part of the DGP additional shifts in the nonstationary part do not exacerbate the tests' small sample power.

5 Empirical Applications

From the previous simulation results we can clearly conclude that ignoring level shifts distorts the inference about the cointegration rank of a system of variables. To be precise, it is more difficult to find the true cointegrating rank. The following two empirical examples demonstrate how these findings may affect applied work. First, I refer to a cointegration analysis of the Polish crawling peg system performed by Trenkler (2000). Furthermore, a German money demand system described by Lütkepohl & Wolters (1998) is used.

5.1 Polish Crawling Peg System

Trenkler (2000) has analysed the relationship between the Polish zloty and the basket currencies within the crawling peg system Poland established on October 15, 1991. In this system the zloty is tied to a currency basket and devalued with a preannounced monthly rate (rate of crawl). The currencies comprising the basket were the U.S. dollar, German mark, British pound, French franc and Swiss franc. Assuming that the monetary

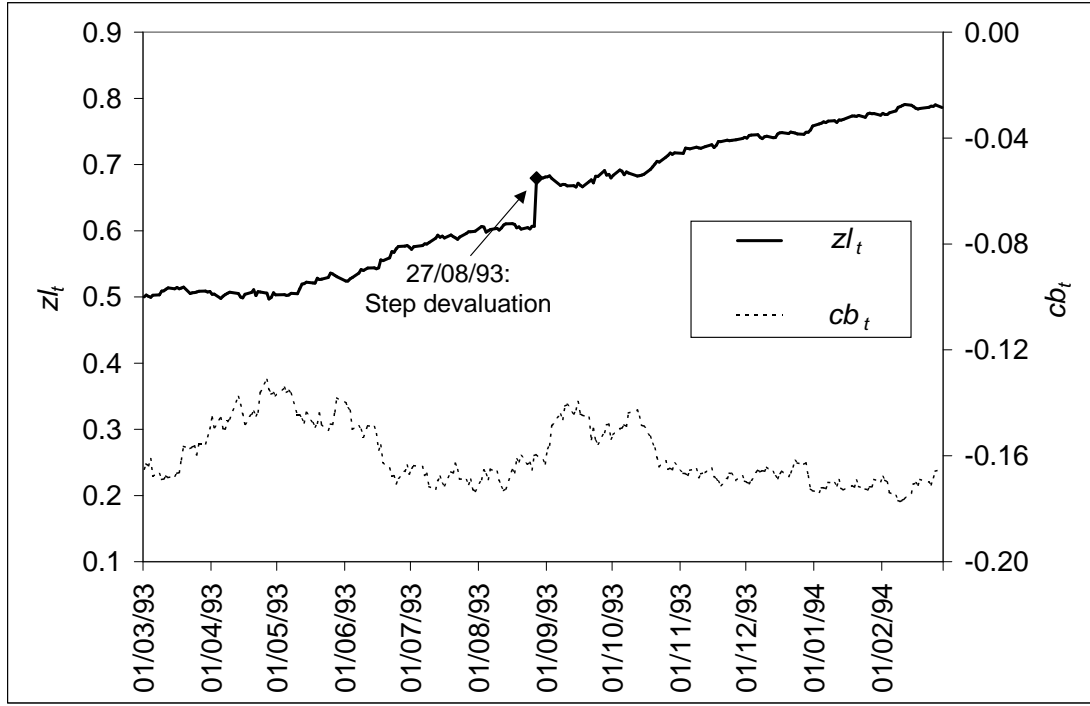


Figure 11. Logarithm of Polish zloty-U.S. dollar rate (zl_t) and U.S. dollar per unit of currency basket (cb_t) between March 1, 1993 - February 28, 1994.

authorities can keep the actual zloty rate in line with the theoretical rate determined by the crawling peg, Trenkler (2000) derives a long-run relationship between the logarithms of the zloty - U.S. dollar rate (zl_t) and the logarithm of the basket currencies denominated in U.S. dollar (gm_t, bp_t, ff_t, sf_t) and a long-run relationship between the logarithms of the zloty - U.S. dollar rate (zl_t) and the logarithm of the basket's value denominated in U.S. dollar (cb_t). Hence, he considers a five-dimensional model consisting of the single currencies and a bivariate model combining the zloty - dollar rate and the basket's value. The former model is named currency model the latter one basket model. Trenkler (2000) tests for these relationships within the cointegration framework using daily data from March 1, 1993 to February 28, 1994. The number of observation is 249. However, this period includes a discrete step devaluation of the Polish zloty by 7.4 % against the basket on August 27, 1993 (see Figure 11). Thus, the shift occurs in the 125th period ($t_1 = 0.5$). This level shift is modelled by a shift dummy within a VAR system of order $p = 3$ consisting of a constant, a linear trend, and $zl_t, gm_t, bp_t, ff_t, sf_t$ (currency model) and zl_t, cb_t (basket model) respectively. By introducing the dummy variable he allows for a shift in all time series. Actually, the break only occurs in zl_t . Therefore, one may expect that the estimated coefficients of the dummy variable related to the other time series are very small. Applying cointegration tests suggested by Saikkonen & Lütkepohl (2000a) (LR_{SL}) and by Johansen, Mosconi & Nielsen (2000) (LR_{JMN}) which take the level shift

Table 7. Cointegration Tests for Basket Model (zl_t, cb_t)

| H_0 | LR_{JMN} | 5% c.v. | LR_{Joh} | 5% c.v. | LR_{SL} | $LR_{S\&L}$ | 5% c.v. |
|-----------|------------|---------|------------|---------|-----------|-------------|---------|
| $r_0 = 0$ | 68.26** | 30.49 | 14.65 | 25.47 | 23.46** | 6.45 | 15.92 |
| $r_0 = 1$ | 10.46 | 15.26 | 3.33 | 12.39 | 2.81 | 2.06 | 6.83 |

Note: ** denotes significance at a 1% level. Critical values (c.v.) for LR_{SL} and $LR_{S\&L}$ are the same and taken from Lütkepohl & Saikkonen (2000). The critical values for LR_{JMN} are simulated with the program DisCo (Johansen & Nielsen 1993). Those for LR_{Joh} are taken from Johansen (1995).

Table 8. Cointegration Tests for Currency Model ($zl_t, gm_t, bp_t, ff_t, sf_t$)

| H_0 | LR_{JMN} | 5% c.v. | LR_{Joh} | 5% c.v. | LR_{SL} | $LR_{S\&L}$ | 5% c.v. |
|-----------|------------|---------|------------|---------|-----------|-------------|---------|
| $r_0 = 0$ | 233.69** | 98.53 | 64.01 | 86.96 | 125.15** | 39.72 | 65.69 |
| $r_0 = 1$ | 49.45 | 72.15 | 42.00 | 62.61 | 28.74 | 27.76 | 45.12 |
| $r_0 = 2$ | 32.23 | 49.38 | 26.24 | 42.20 | 10.63 | 11.26 | 28.47 |
| $r_0 = 3$ | 17.29 | 30.49 | 11.94 | 25.47 | 5.67 | 5.50 | 15.92 |
| $r_0 = 4$ | 4.39 | 15.26 | 4.27 | 12.39 | 1.65 | 2.44 | 6.83 |

Note: ** denotes significance at a 1% level respectively. Critical values (c.v.) for LR_{SL} and $LR_{S\&L}$ are the same and taken from Lütkepohl & Saikkonen (2000). The critical values for LR_{JMN} are simulated with the program DisCo (Johansen & Nielsen 1993). Those for LR_{Joh} can be found in Johansen (1995).

into account, Trenkler (2000) finds a cointegrating rank of one for both models confirming the existence of the postulated long-run relationships. These tests are based on (2.1) and (2.9) and work similarly compared to the standard tests. For the Johansen test, y_{t-1}^+ as defined in (2.9) is used instead of the definition in (2.11). Otherwise, the procedure applies as described in Section 2. However, one has to keep in mind that the distribution of the test statistic differs from the one for (2.13) and depends on the timing of the shift. Critical values can be calculated by using the simulation program DisCo (see Johansen & Nielsen 1993) or by a response surface given in Johansen et al. (2000). For the Saikkonen-Lütkepohl proposal, (2.9) is used to obtain first-stage estimators. Then all deterministic terms in (2.1) including the level shift are estimated by a GLS procedure. The time series y_t is adjusted by these deterministic terms and an LR type test is performed on the adjusted series $\hat{x}_t = y_t - \hat{\mu}_0 - \hat{\mu}_1 t - \hat{\delta} d_t$. The limiting distribution of the resulting test statistic is the same as in the case of no level shifts. Details of the test are described in Saikkonen & Lütkepohl (2000a).

Table 9. Estimation of Scaled Shift Parameters for Polish Crawling Peg System, March 1, 1993 - February 28, 1994, $T = 249$

| Time Series | $\hat{\delta}$ | $\hat{\sigma}$ | $\hat{\delta}_{100}^*$ | $\hat{\delta}_{200}^*$ |
|--|----------------|----------------|------------------------|------------------------|
| Basket Model (z_{l_t}, cb_t) | | | | |
| $e_{\Zl | 0.07661 | 0.002012 | 24.13 | 34.12 |
| Currency Model ($z_{l_t}, gm_t, bp_t, ff_t, sf_t$) | | | | |
| $e_{\Zl | 0.07609 | 0.001665 | 28.96 | 40.95 |

Trenkler (2000) has also applied the standard cointegration tests LR_{Joh} and $LR_{S\&L}$ analysed in the foregoing section and found no cointegration. The results are collected in Tables 7 and 8.

In contrast to the simulation study we do not know the true cointegrating rank of the systems under study. Therefore, we cannot asses the performance of the tests as we have done in the previous section. Nevertheless, we have the results of the corresponding cointegration tests accommodating a level shift. If we assume that these tests find the true cointegrating rank, then one may take them as a reference. To gauge the test results in an useful way it is also necessary to estimate the magnitude of the level shift. The GLS procedure used by Saikkonen & Lütkepohl (2000a) provides an estimator $\hat{\delta}$ for the parameter vector δ as well. In line with the results of LR_{JMN} and LR_{SL} I apply a cointegrating rank of one since I assume that this is the true cointegrating rank. The asymptotic properties of this estimator given in Theorem 3.1 of Saikkonen & Lütkepohl (2000a) state that δ is consistently estimated in direction of β if $T - T_1 \rightarrow \infty$, but is only bounded in probability in the direction of β_{\perp} . Therefore, one has to use some caution in interpreting the estimated coefficient.

The estimation results for $\hat{\delta}$ for z_{l_t} are reported in Table 9. The estimated shift parameters for the other series are equal to or less than 0.2 (currency model) and equal to 0.5 (basket model) if normalized for $T = 100$. Therefore, these results are not shown in detail in Table 9. The shift estimates are roughly equal for both models and close to the actual devaluation rate of 7.4% or 0.074.

These estimates, however, have to be related to the sample size and the standard deviation of the error term in the respective component according to the scaling problem discussed in the foregoing sections. We can estimate the variance-covariance matrix of the error terms by using the residuals from the reduced rank regression of (2.9) and thereby extract estimators of the standard deviations ($\hat{\sigma}$). In order to simplify a comparison with

the results of the Monte Carlo study the estimated magnitudes $\hat{\delta}$ are scaled, or we may say normalized, with respect to the sample sizes $T = 100$ and $T = 200$ used in the Monte Carlo study. Including the estimated standard deviations in (4.5) I obtain for the current sample size $T = 249$ the normalized shift magnitudes as

$$\hat{\delta}_i^{(100)} = \frac{\sqrt{(100/249)} \hat{\delta}_i}{\hat{\sigma}_i} \quad \text{and} \quad \hat{\delta}_i^{(200)} = \frac{\sqrt{(200/249)} \hat{\delta}_i}{\hat{\sigma}_i}$$

where the index i corresponds to the estimated parameter values of the respective time series in the basket and currency models. The resulting quantities for zl_t are shown in Table 9.

Since I have allowed the level shift to enter the long-run relationship according to (2.9) we may interpret it as a level shift in the stationary part with respect to the simulation study presented above, i.e. the shift is assumed to be not orthogonal to the cointegration space. Then, taking the large magnitude of the level shift in the Polish zloty series into account we should expect that the tests suffer from a high power loss since these values are far above the magnitudes analysed in the foregoing section. Figure 4 shows for comparable bivariate DGPs with $T = 200$, a cointegrating rank $r = 1$ and strong innovation correlation that the power is relatively low in case of the largest magnitude I have analyzed. Although a shift has occurred in both components of these DGPs it is known from the results for $T = 100$ that the outcomes for a break in the stationary part only are similar to breaks in all components.⁷ It was also stated that the results for four-dimensional systems are equivalent to the ones for bivariate models if a cointegrating of $r = 1$ is considered. In this respect it is not surprising that the standard tests cannot find a cointegration relationship neither in the five-dimensional currency model nor the bivariate basket model although the tests accommodating the level shift reject the null hypothesis of no cointegration at a 1% level.

5.2 A German Macroeconomic System

Lütkepohl & Wolters (1998) aim to analyse the channels of the German monetary policy by means of a small model for the monetary sector that comprises M3, GNP, an inflation rate, an interest-rate spread variable measuring the difference between the average bond rate and the own rate of M3, and import-price inflation. The center of the model is a demand relation for M3, the target variable for the monetary policy of the Deutsche Bundesbank. Within this relation GNP is used as a variable for the transaction volume,

⁷In the current example a level shift occurs only in one time series and, as mentioned above, this level shift is assumed to correspond to breaks in the stationary part of the DGPs used in the Monte Carlo study.

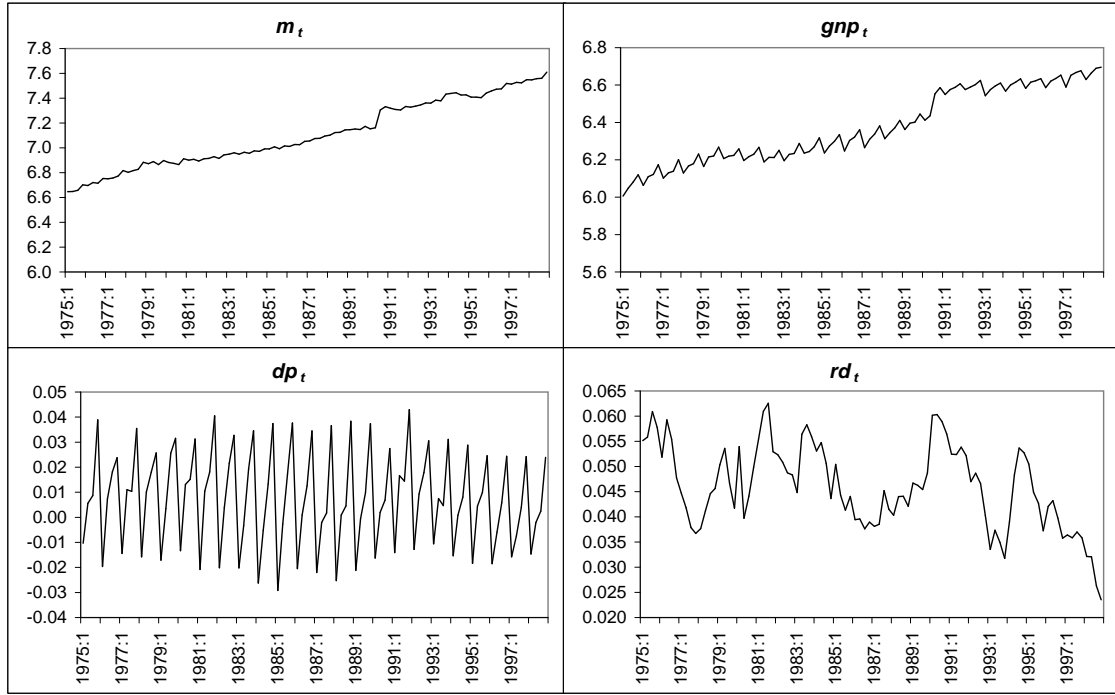


Figure 12. Time series of German macroeconomic data, 1975:1 - 1998:4.

the inflation rate and the interest-rate spread are included to represent opportunity costs of money holding and the import-price inflation accounts for the openness of the German economy. Using the same data the unit root analysis conducted by Wolters, Teräsvirta & Lütkepohl (1998) has found that the interest-rate spread and the import-price inflation are stationary, whereas the other variables are $I(1)$.

Like Saikkonen & Lütkepohl (2000a) I do not analyse this German money demand system to derive detailed economic implications as Lütkepohl & Wolters (1998) and Wolters et al. (1998) have done it but rather I am using the system to illustrate certain econometric aspects that can be important for empirical work. Saikkonen & Lütkepohl (2000a) have used the money demand system to apply their cointegration tests accommodating level shifts to an actual economic system. Furthermore, they demonstrate that it is worthwhile to apply cointegration tests allowing for level shifts to the whole sample period instead of applying standard tests to a data set ending before the occurrence of the level shift. The reason is that in the latter case one may find a lower cointegrating rank than the expected one due to reduced power resulting from using a shorter sample period. I am interested, however, in studying the effects of ignoring level shifts in the data set as mentioned before.

According to the cointegration analysis of the same system in Saikkonen & Lütkepohl (2000a) I test for the number of cointegration relations between the nonstationary vari-

Table 10. Cointegration Tests for German Macroeconomic Data ($m_t, gnp_t, \Delta p_t$), 1975:1 - 1998:4, $T = 96$

| H_0 | LR_{JMN} | 5% c.v. | LR_{Joh} | 5% c.v. | LR_{SL} | $LR_{S\&L}$ | 5% c.v. |
|---------|------------|---------|------------|---------|-----------|-------------|---------|
| $r_0=0$ | 91.34** | 49.63 | 84.59** | 42.20 | 62.50** | 45.47** | 28.47 |
| $r_0=1$ | 18.32 | 30.98 | 15.49 | 25.47 | 5.16 | 6.98 | 15.92 |
| $r_0=2$ | 6.54 | 15.28 | 2.87 | 12.33 | 1.43 | 2.39 | 6.83 |

Note: ** denotes significance at a 1% level. Critical values (c.v.) for LR_{SL} and $LR_{S\&L}$ are the same and taken from Lütkepohl & Saikkonen (2000). The critical values for LR_{JMN} are simulated with the program DisCo (Johansen & Nielsen 1993). Those ones for LR_{Joh} are taken from Johansen (1995).

ables M3, GNP, and the inflation rate. In a second step the (stationary) interest-rate spread is added in order to analyse whether a further cointegration relation is detected the cointegration tests.

As in Lütkepohl & Wolters (1998) and Saikkonen & Lütkepohl (2000a) I use quarterly, seasonally unadjusted data, but extend the original period 1975:1 -1996:4 to the fourth quarter in 1998 which is the last period before the introduction of the euro. This period has also been analysed by Lütkepohl & Wolters (2001). They consider a similar model as Lütkepohl & Wolters (1998). The main difference between these two studies is that Lütkepohl & Wolters (2001) use separate long-term and short-term interest rates instead of an interest rate spread in order to examine the monetary policy of the Bundesbank directly.

The initial period 1975:1 represents the beginning of the Bundesbank's policy of monetary targeting. To be precise, the following variables are used: m_t is the logarithm of real M3, gnp_t is the logarithm of real GNP, Δp_t represents the inflation rate as the first difference of the logarithm of the GNP deflator, and rd_t is the difference between the average bond rate and the own rate of M3 as described above. The data sources are given in the Appendix A. The variables are displayed in Figure 12. Obviously, m_t and gnp_t exhibit level shifts in 1990:3 due to the German reunification. In July 1990 the West and East German time series have been aggregated. Saikkonen & Lütkepohl (2000a) model these level shifts within a VAR system by introducing a shift dummy variable set to zero till the second quarter of 1990 and one afterwards and an impulse dummy that is one in 1990:3 and zero elsewhere. Furthermore, a linear trend, a constant and seasonal dummy variables are added to meet the obvious properties of the time series. Saikkonen & Lütkepohl (2000a) use a VAR order of two which is also suggested for the extended sample period by

Table 11. Cointegration Tests for German Macroeconomic Data ($m_t, gnp_t, \Delta p_t, rd_t$), 1975:1 - 1998:4, $T = 96$

| H_0 | LR_{JMN} | 5% c.v. | LR_{Joh} | 5% c.v. | LR_{SL} | $LR_{S\&L}$ | 5% c.v. |
|-----------|------------|---------|------------|---------|-----------|-------------|---------|
| $r_0 = 0$ | 127.93** | 72.30 | 107.12** | 62.61 | 87.33** | 77.13** | 45.13 |
| $r_0 = 1$ | 54.56* | 49.63 | 36.12 | 42.20 | 32.53* | 24.24 | 28.47 |
| $r_0 = 2$ | 20.26 | 30.98 | 15.72 | 25.47 | 5.87 | 7.80 | 15.92 |
| $r_0 = 3$ | 9.57 | 15.28 | 3.42 | 12.33 | 0.57 | 2.02 | 6.83 |

Note: ** and * denote significance at a 1% level and 5% level respectively. Critical values (c.v.) for LR_{SL} and $LR_{S\&L}$ are the same and taken from Lütkepohl & Saikkonen (2000). The critical values for LR_{JMN} are simulated with the program DisCo (Johansen & Nielsen 1993). Those ones for LR_{Joh} are taken from Johansen (1995).

the Schwarz (SC) and Hannan-Quinn (HQ) criteria (see Lütkepohl (1991, Chapter 4) for definitions). Therefore, I apply this order as well. Thus, the tests allowing for level shifts are based on the VAR(2) model including all mentioned deterministic terms whereas the standard tests refer to a VAR(2) model without the shift dummy and impulse dummy variable. Furthermore, performing a unit root analysis for the extended sample I obtain the same results as Wolters et al. (1998), i.e. m_t , gnp_t , and Δp_t are $I(1)$ whereas rd_t is stationary. The outcome for m_t , gnp_t , and Δp_t is in line with the one in Lütkepohl & Wolters (2001) who also include these time series.

The results of the cointegration tests are summarized in Tables 10 and 11. Using the tests that accommodate the level shift a cointegrating rank of one for the three-dimensional system ($m_t, gnp_t, \Delta p_t$) is found. These tests also detect two cointegration relations in the larger system ($m_t, gnp_t, \Delta p_t, rd_t$) including the interest rate spread variable in accordance with the stationarity of this variable. Applying the standard cointegration tests to the smaller system one can see that ignoring the level shift does not change the inference about the cointegrating rank although the values of the test statistics are lower for the null hypothesis $H_0(0) : r_0 = 0$. However, if $LR_{S\&L}$ and LR_{Joh} are applied to the four-dimensional system including the interest-rate spread variable only a cointegrating rank of one is found. These tests could not reject the null of one cointegration relation neither at a 5% nor at a 10% significance level.

To evaluate these findings the scaled magnitudes for the level shifts are estimated. The GLS estimators $\hat{\delta}$ are computed as described for the Polish crawling peg system. The results for $\hat{\delta}$, the standard deviation of the respective error term $\hat{\sigma}$, and the normalized parameters vectors $\hat{\delta}_i^{(100)}$ and $\hat{\delta}_i^{(200)}$ are reported in Table 12. Although the estimators for

Table 12. Estimation of Scaled Shift Parameters for German Macroeconomic Data, 1975:1 - 1998:4, $T = 96$

| Time series | $\hat{\delta}$ | $\hat{\sigma}$ | $\hat{\delta}^{(100)}$ | $\hat{\delta}^{(200)}$ |
|--|----------------|----------------|------------------------|------------------------|
| Three-dimensional system $(m_t, gnp_t, \Delta p_t)$ | | | | |
| m_t | 0.13744 | 0.00973 | 14.42 | 20.39 |
| gnp_t | 0.10465 | 0.01302 | 8.20 | 11.60 |
| Δp_t | -0.00309 | 0.00662 | -0.48 | -0.67 |
| Four-dimensional system $(m_t, gnp_t, \Delta p_t, rd_t)$ | | | | |
| m_t | 0.13812 | 0.00941 | 14.98 | 21.18 |
| gnp_t | 0.10539 | 0.01300 | 8.28 | 11.71 |
| Δp_t | -0.00339 | 0.00658 | -0.53 | -0.74 |
| rd_t | -0.00224 | 0.00447 | -0.51 | -0.72 |

the normalized level shift in m_t differ depending on the two systems, they are approximately of the same order and similar to the largest values for the shift magnitudes I have analysed in the simulation study. In contrast, the shift in the GNP series is of a medium size. Furthermore, since the shifts in dp_t and r_t are very small one can assume that a shift only occurs in two (m_t, gnp_t) of the three or four time series respectively.

Referring to the simulation results for bivariate and four-dimensional DGPs with $r = 1$ and strong error term correlation ($\theta = 0.8$) one can see that the standard cointegration tests have some power if scaled shift magnitudes (for $T = 100$) between 5 and 10 are considered (compare Figure 1)⁸. This may explain why a cointegration rank of one is still found in the small system. However, we have to keep in mind that the shift in m_t is of larger size.

We have learnt from the discussion of the simulation results related to four-dimensional DGPs with $r = 2$ that there is some probability to find one of two cointegration relations (compare also Figures 5-8). But it is very unlikely to detect a second relationship. In line with these observations $LR_{S\&L}$ and LR_{Joh} find only a cointegrating rank of one for the four-dimensional system, i.e. neither of the standard cointegration tests is able to infer the cointegrating rank of two which is found by the tests taking account of level shifts.

Moreover, we can conclude from Tables 10 and 11 that both the Johansen and the Saikkonen-Lütkepohl proposals are affected in a similar way when the level shift is ignored. This observation is in line with the corresponding findings of the simulation study

⁸The findings for four-dimensional DGPs with $r = 1$ are similar.

for the four-dimensional DGPs. The three-dimensional model within the illustration on the German money demand system can be regarded as an empirical example for which neglecting a level shift may lead to similar inference about the cointegrating rank as in models where the shift is taken into account explicitly.

6 Summary

In this study I have analysed the effects of ignoring level shifts in the data generating process on systems cointegration tests. This analysis is conducted using Monte Carlo studies and reviewing two empirical examples. I have considered two groups of Likelihood Ratio tests based on procedures suggested by Johansen (1991, 1992, 1994, 1995) and Saikkonen & Lütkepohl (2000b). These procedures specify test versions allowing for a mean term only and versions additionally including a linear trend.

With respect to the Monte Carlo results I can say that ignoring level shifts leads to size distortions in such a way that the tests' size approaches zero for increasing values of the shift magnitude. Furthermore, level shifts in the stationary part of the process resulting in shifts in the cointegration relations are much more important than shifts in the nonstationary part which leads to unrestricted impulses in the VECM representation only. To be more precise, the former kind of shifts produces significant size reductions for lower values of the shift magnitude. Nevertheless, the sizes approach zero for both types of breaks as the shift size increases.

For moderate shift magnitudes the Saikkonen & Lütkepohl tests are sometimes less distorted than the Johansen proposals. On the other hand, the different allowance made for the deterministic terms, the number of components exhibiting a shift and other characteristics of the DGPs like the dimension or the error term correlation do not cause important or systematic differences in terms of the size distortion. The same can be said about the sample size if the magnitudes of the shifts are scaled by the sample size.

Obviously, the power of all tests is clearly reduced for increasing shift magnitudes. In case of four-dimensional DGPs with $r = 2$ the tests still have some power (30 – 40%) if the null hypothesis $r_0 = 0$ is tested even if the size of the break is large. For the bivariate DGPs we also observe that the tests allowing for a linear trend lose less power than the tests assuming no trend resulting in an absolute power advantage for the former group of tests. With respect to the Johansen and Saikkonen & Lütkepohl test proposals no important differences in their relative performance are observed. In general, the values of the other parameters characterizing the DGP, like the autoregressive eigenvalues or the innovation correlation, seem to have just an impact on the level of the tests' small sample

power instead of determining how the power is reduced by the introduction of a level shift.

The timing of the level shift is in general of minor importance. However, shifts occurring at the very beginning of the sample do not generate reductions in the size in small samples or power of the Johansen test. In contrast, for shifts at the end of the sample I do not observe special effects neither for the Johansen nor for the Saikkonen-Lütkepohl tests.

The analysis has been restricted to a single level shift at one point in time that may affect all or only some of the components of the multivariate time series. It is expected that the existence of more than one level shift distorts the tests' inference even more, i.e. it is even more difficult to detect the true cointegrating rank if more than one shift is present.

We have to keep in mind that no asymptotical results on the effects of ignoring level shifts on systems cointegration tests are available yet. Although, the size in small samples approaches zero for increasing shift magnitudes, the test still have some small sample power in different situations. Hence, one should be very careful in drawing conclusion about the tests' asymptotic properties. In this respect, the asymptotic results related to corresponding unit root tests may only be guidelines.

The review of the empirical applications illustrates the main results of the Monte Carlo study. Applying cointegration tests that ignore level shifts to systems that in fact exhibit such shifts leads to a distorted inference about the cointegrating rank. To be precise, a smaller cointegration rank than the one suggested by tests which accommodate the shifts is found in general. However, in accordance with the simulation results there may occur situations when ignoring level shifts does not affect the tests' inference about the cointegration rank. Moreover, as the small sample power analysis the application also demonstrates that both $LR_{S\&L}$ and LR_{Joh} are affected in a similar way when shifts are ignored.

Hence, the evidence from the Monte Carlo study and the empirical applications strongly suggests to use systems cointegration tests that accommodate level shifts because the distortionary effects of such shifts on the small sample properties of the standard tests not allowing for shifts cannot be ignored. Systems cointegration tests that take level shifts into account have been proposed by Johansen et al. (2000) and Saikkonen & Lütkepohl (2000a). Their performance has been studied by Lütkepohl, Saikkonen & Trenkler (2000) and Trenkler (2002) by means of local power and small sample analyses. In general, the performance of the procedures allowing and not allowing for level shifts is rather similar provided that both test groups are applied in appropriate situations. Therefore, tests

taking level shifts into consideration are recommended for applied work if such shifts are in fact present.

References

- Campos, J., Ericsson, N. & Hendry, D. F. (1996). Cointegration tests in the presence of structural breaks, *Journal of Econometrics* **70**: 187–220.
- Doornik, J., Hendry, D. F. & Nielsen, B. (1998). Inference in cointegrating models: UK M1 revisited, *Journal of Econometric Surveys* **12**: 533–571.
- Engle, R. F. & Granger, C. W. J. (1987). Cointegration and error correction: Representation, estimation and testing, *Econometrica* **55**: 251–276.
- Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification, *Journal of Econometrics* **16**: 121–131.
- Gregory, A. W. & Hansen, B. E. (1996). Residual-based tests for cointegration in models with regime shifts, *Journal of Econometrics* **70**: 99–126.
- Gregory, A. W., Nason, J. M. & Watt, D. G. (1996). Testing for structural breaks in cointegrated relationships, *Journal of Econometrics* **71**: 321–341.
- Hubrich, K., Lütkepohl, H. & Saikkonen, P. (2001). A review of systems cointegration tests, *Econometric Reviews* **20**: 247–318.
- Inoue, A. (1999). Tests of cointegrating rank with a trend-break, *Journal of Econometrics* **90**: 215–237.
- Johansen, S. (1988). Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control* **12**: 231–254.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models, *Econometrica* **59**: 1551–1580.
- Johansen, S. (1992). Determination of cointegration rank in the presence of a linear trend, *Oxford Bulletin of Economics and Statistics* **54**: 383–397.
- Johansen, S. (1994). The role of the constant and linear terms in cointegration analysis of nonstationary time series, *Econometric Reviews* **13**: 205–231.

- Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.
- Johansen, S. & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration - with applications to the demand for money, *Oxford Bulletin of Economics and Statistics* **52**: 169–210.
- Johansen, S., Mosconi, R. & Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend, *Econometrics Journal* **3**: 216–249.
- Johansen, S. & Nielsen, B. (1993). Manual for the simulation program DisCo. Institute of Mathematical Statistics, University of Copenhagen.
- Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.
- Lütkepohl, H. & Saikkonen, P. (2000). Testing for the cointegrating rank of a VAR process with a time trend, *Journal of Econometrics* **95**: 177–198.
- Lütkepohl, H., Saikkonen, P. & Trenkler, C. (2000). Comparison of tests for the cointegrating rank of a VAR process with a structural shift, *Discussion Paper 10*, SFB 373, Humboldt-Universität zu Berlin.
- Lütkepohl, H., Saikkonen, P. & Trenkler, C. (2001). Maximum eigenvalue versus trace tests for the cointegrating rank of a VAR process, *Econometrics Journal* **4**: 287–310.
- Lütkepohl, H. & Wolters, J. (1998). A money demand system for M3 in the unified germany, *Empirical Economics* **23**: 371–386.
- Lütkepohl, H. & Wolters, J. (2001). The transmission of German monetary policy in the pre-euro period, *Discussion Paper 87*, SFB 373, Humboldt-Universität zu Berlin.
- Maddala, G. & Kim, I.-M. (1998). *Unit Roots, Cointegration, and Structural Change*, Cambridge University Press.
- Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis, *Econometrica* **57**: 1361–1401.
- Perron, P. (1990). Testing for a unit root in a time series with a changing mean, *Journal of Business & Economic Statistics* **8**: 301–320.

- Perron, P. & Campbell, J. Y. (1993). A note on Johansen's cointegration procedure when trends are present, *Empirical Economics* **18**: 777–789.
- Saikkonen, P. & Lütkepohl, H. (2000a). Testing for the cointegrating rank of a VAR process with structural shifts, *Journal of Business & Economic Statistics*, **18**: 451–464.
- Saikkonen, P. & Lütkepohl, H. (2000b). Trend adjustment prior to testing for the cointegrating rank of a VAR process, *Journal of Time Series Analysis* **21**: 435–456.
- Toda, H. Y. (1994). Finite sample properties of likelihood ratio tests for cointegrating ranks when linear trends are present, *The Review of Econometrics and Statistics* **76**: 66–79.
- Toda, H. Y. (1995). Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions, *Econometric Theory* **11**: 1015–1032.
- Trenkler, C. (2000). The polish crawling peg system: A cointegration analysis, *Discussion Paper 71*, SFB 373, Humboldt-Universität zu Berlin.
- Trenkler, C. (2002). *Testing for the Cointegrating Rank in the Presence of Level Shifts*, Shaker Verlag, Berlin.
- Wolters, J., Teräsvirta, T. & Lütkepohl, H. (1998). Modelling the demand for M3 in the unified Germany, *Review of Economics and Statistics* **80**: 399–409.
- Yin, Y. & Maddala, G. (1997). The effects of different types of outliers on unit root tests, in T. Fomby & R. Hill (eds), *Advance in Econometrics*, Vol. 13, JAI Press, Greenwich, Conn.